Adaptive quantile estimation in deconvolution with unknown error distribution

Itai Dattner[*], **Markus Reiß**[*] **and Mathias Trabs**[*]

EURANDOM
Department of Mathematics and Computer Science
Eindhoven University of Technology
P.O. Box 513
5600 MB Eindhoven, The Netherlands
e-mail: dattner@eurandom.tue.nl

Institut für Mathematik
Humboldt-Universität zu Berlin
Unter den Linden 6
10099 Berlin, Germany
e-mail: mreiss@math.hu-berlin.de
trabs@math.hu-berlin.de

Abstract: We study the problem of quantile estimation in deconvolution with ordinary smooth error distributions. In particular, we focus on the more realistic setup of unknown error distributions. We develop a minimax optimal procedure and construct an adaptive estimation method under natural conditions on the densities. As a side result we obtain minimax optimal rates for the plug-in estimation of distribution functions with unknown error distributions.

1. Introduction

Let $X_1, \ldots, X_n$ be a sequence of independent identically distributed random variables with a common Lebesgue density $f : \mathbb{R} \to \mathbb{R}$. Suppose that we observe random variables $Y_1, \ldots, Y_n$, $n \in \mathbb{N}$, given by

$$Y_j = X_j + \varepsilon_j, \quad j = 1, \ldots, n, \tag{1}$$

where ε_j are i.i.d. random variables, independent of (X_j) with the Lebesgue density f_ε. For $\tau \in (0, 1)$ the objective is to estimate the τ-quantile of the population X given by

$$q_\tau = \arg\min_{\eta \in \mathbb{R}} \int_{-\infty}^{\infty} \rho_\tau(x - \eta) f(x) \, dx \quad \text{with} \quad \rho_\tau(x) = x(\tau - \mathbb{1}_{\{x \leqslant 0\}}) \tag{2}$$

from the observations $Y_1, \ldots, Y_n$. Note that with the above definition of ρ_τ, a finite first moment of X_j would be necessary. To avoid this assumption, $\rho_\tau(x, \eta) = \tau\big((x - \eta)_+ - x_+\big) + (1 - \tau)\big((x - \eta)_- - x_-\big)$, for $x, \eta \in \mathbb{R}$, can be used instead of $\rho_\tau(x - \eta)$ in (2).

Assuming that the distribution of the measurement error is completely known, Carroll and Hall (1988) have constructed a kernel density estimator based on the empirical characteristic function $\varphi_n(u) := \frac{1}{n} \sum_{j=1}^{n} e^{iuY_j}, u \in \mathbb{R}$. In practice, however, the distribution of the measurement error is usually not known. Instead, we assume that we have at hand a sample from f_ε given by

$$\varepsilon_1^*, \ldots, \varepsilon_m^*, \qquad m \in \mathbb{N}. \tag{3}$$

Motivated from applications, we will not assume that the observations (ε_k^*) are independent from (Y_j). In particular, our procedure applies to the experiment setup of repeated measurements, which we will discuss below.

We define the Fourier transform of $g \in L^1(\mathbb{R}) \cup L^2(\mathbb{R})$ by $\mathcal{F} g(u) := \int_{\mathbb{R}} e^{iux} g(x) \, \mathrm{d}x$ for $u \in \mathbb{R}$. Based on the classical kernel estimator, Neumann (1997) has proposed the following density estimator for unknown error distributions

$$\widetilde{f}_b(x) := \mathcal{F}^{-1} \left[\frac{\varphi_n(u) \varphi_K(bu)}{\varphi_{\varepsilon,m}(u)} \right](x), \quad x \in \mathbb{R}, \tag{4}$$

where φ_K is the Fourier transform of the kernel K, $b > 0$ is the bandwidth and the characteristic function of the error distribution φ_ε is estimated by its empirical counterpart $\varphi_{\varepsilon,m}(u) := \frac{1}{m} \sum_{k=1}^m e^{iu\varepsilon_k^*}, u \in \mathbb{R}$. Obviously, $\widetilde{f}_b$ depends on the sample sizes n and m which is suppressed in the notation. Applying a plug-in approach, our estimator for the quantile q_τ is then given by the quantile of the estimated density

$$\widetilde{q}_{\tau,b} = \operatorname*{arg\,min}_{\eta \in \mathbb{R}} \widetilde{M}_b(\eta) \quad \text{with} \quad \widetilde{M}_b(\eta) := \int_{-\infty}^{\infty} \rho_\tau(x - \eta) \widetilde{f}_b(x) \, \mathrm{d}x. \tag{5}$$

Throughout the text we assume that $\varphi_\varepsilon(u) \neq 0, u \in \mathbb{R}$. In particular, in this work we are interested in error distributions for which the characteristic function decays polynomially in its tails. As shown by Fan (1991), these so-called ordinary smooth errors lead to mildly ill-posed estimation problems.

Although the literature on deconvolution problems is extensive, the problem of adaptive deconvolution with unknown measurement errors was only recently addressed. We refer to Comte and Lacour (2011); Johannes and Schwarz (2010) and Kappus (2012) for adaptive deconvolution of densities with unknown error distributions in the model selection framework. Minimax results and other properties for non-adaptive methods are given in Neumann (1997), Meister (2004), Neumann (2007), Delaigle, Hall and Meister (2008), Johannes (2009) among others. To the best of our knowledge, the problem of quantile estimation in deconvolution was considered only in Hall and Lahiri (2008). They constructed a non-adaptive quantile estimator for the case of known error distributions. Unlike the estimator (2), their estimator is based on directly inverting the distribution function estimator. Indeed, following the classical M-estimator analysis, the error of the quantile estimator (refeq:quantile) is directly related to that of the distribution function estimator (cf. the error representation (8) below). However, the analysis of the later was not clear until recently.

Fan (1991) constructed an estimator for the distribution function by integrating the density deconvolution estimator. In order to perform an exact analysis of its variance a truncation of the integral was required in the estimation procedure. This resulted in a non-optimal estimation method for the case of ordinary smooth errors. Trying to circumvent this problem, Hall and Lahiri (2008) as well as Dattner, Goldenshluger and Juditsky (2011) constructed a distribution deconvolution estimator based on a direct inversion formula for distribution functions. Applying the Fourier multiplier approach by Nickl and Reiß (2012), Söhl and Trabs (2012) have shown that the integrated density estimator can estimate the distribution function with $\sqrt{n}$-rate under suitable conditions. Still, they do not cover all cases where a parametric rate should be expected. So even with a known error distribution, a rigorous answer to the following question was left open: can the canonical plug-in estimator be used for minimax optimal distribution function estimation? Can quantiles be estimated optimally by the plug-in approach, too?

To show that the answer to these question is 'yes', we combine an exact analysis like the one by Dattner, Goldenshluger and Juditsky (2011) together with abstract Fourier multiplier theory as in Söhl and Trabs (2012). This combination yields the required minimax optimal results for the distribution function and for quantiles. This closes the gap reported by Fan (1991) for distribution deconvolution in the ordinary smooth case. Moreover, we show that the minimax optimal rates still hold if the error distribution is unknown and has to be estimated, which is mathematically challenging. Since the deconvolution operator $\mathcal{F}^{-1}[1/\varphi_\varepsilon]$ is not observable, we have to study the

estimated counterpart $\mathcal{F}^{-1}[\varphi_K(bu)/\varphi_{\varepsilon,m}(u)]$. As a random Fourier multiplier it preserves the mapping properties of the deterministic $\mathcal{F}^{-1}[1/\varphi_\varepsilon]$, but its operator norm turns out to be (slightly) larger.

We can summarize our main contributions as follows:

(i) The case of unknown error distributions is studied and minimax rates of convergence of the estimator $\widetilde{q}_{\tau,b}$ are established. Unlike previous studies, we do not assume that the observations (ε_k^*) are independent from (Y_j). Moreover, all results carry over to the case of deconvolution of distributions and quantiles under a known error distribution.

(ii) Using the Lepskiĭ (1990) method, an adaptive version of the estimator $\widetilde{q}_{\tau,b}$ is constructed whose convergence rates lose only a logarithmic factor compared to the oracle estimator.

(iii) Compared to previous results, the conditions on the density f are significantly weaker. In a natural way, the rates depend only on the local Hölder smoothness of the density f around the true quantile and the decay rate of φ_ε. Neither tail conditions nor global smoothness conditions on f are needed.

The rest of this paper is organized as follows. In Section 2 we establish the minimax properties of the quantile estimator for both the known and unknown error distribution cases. The adaptive estimation is developed in Section 3. In Section 4 we apply our estimation procedure in simulations and a real data example. The proofs are postponed to the appendix.

2. Convergence rates

Before we start with the error analysis, let us describe the class of densities we are interested in. Denoting $\langle\alpha\rangle$ as the largest integer which is strictly smaller than $\alpha > 0$, we define for some function g and any possibly unbounded interval $I \subseteq \mathbb{R}$ the Hölder norm

$$\|g\|_{C^\alpha(I)} := \sum_{k=0}^{\langle\alpha\rangle} \|g^{(k)}\|_\infty + \sup_{x,y\in I:x\neq y} \frac{|g^{\langle\alpha\rangle}(x) - g^{\langle\alpha\rangle}(y)|}{|x-y|^{\alpha-\langle\alpha\rangle}} \tag{6}$$

and with $R > 0$

$$C^\alpha(I,R) := \big\{g \in C^0(I)\big|\|g\|_{C^\alpha(I)} < R\big\},$$

where $C^0(I)$ is the space of all continuous and bounded functions on the interval I. Let $\mathcal{F}(R)$ denote the set of all probability densities on the real line which are uniformly bounded by some constant $R > 0$. Throughout, we consider for finite positive constants R, r, ζ and the smoothness index $\alpha > 0$ the class

$$\mathcal{C}^\alpha(R,r,\zeta) := \Big\{f \in \mathcal{F}(R)\Big|f \text{ has a } \tau\text{-quantile } q_\tau \in \mathbb{R} \text{ such that}$$

$$f \in C^\alpha([q_\tau - \zeta, q_\tau + \zeta], R) \text{ and } f(q_\tau) \geqslant r\Big\}.$$

Note that the quantile q_τ is unique given the assumption $f(q_\tau) > 0$. Taking the derivative in (5) and restricting to a growing interval, the estimator $\widetilde{q}_{b,\tau}$ can be defined as solution of the estimating equation

$$0 = \widetilde{M}_b'(\eta) = \int_{-\infty}^\eta \widetilde{f}_b(x)\,dx - \tau \quad \text{for some} \quad \eta \in [-U_n, U_n]. \tag{7}$$

Throughout, let (U_n) be a sequence of positive real numbers with $U_n \to \infty$ and which are of the order $\log n$. In view of finite computational time, any algorithm has to restrict to a bounded interval. Since q_τ is fixed for $f \in \mathcal{C}^\alpha(R,r,\zeta)$, the true quantile will be contained in this interval for n sufficiently large. Equation (7) illustrates the relation to the approach by Hall and Lahiri (2008) who invert the distribution function. However, without further assumptions $\widetilde{f}_b$ will not be integrable in general. As a key result in our analysis we state the following lemma. In particular, it will be used to show that the integral in (7) is finite (cf. decomposition (10) below).

In the sequel we use the Landau notation $\mathcal{O}$ and $\mathcal{O}_P$. For two sequences $A_n(\vartheta), B_n(\vartheta)$ depending on a parameter ϑ, $A_n(\vartheta) = \mathcal{O}_P(B_n(\vartheta))$ holds uniformly over a parameter set $\vartheta \in \Theta$ if there is for all $c > 0$ some $C > 0$ such that $\sup_{\vartheta \in \Theta} P_\vartheta(A_n(\vartheta) > CB_n(\vartheta)) < c$. If $A_n(\vartheta)/B_n(\vartheta)$ converges in probability to zero, we write $A_n(\vartheta) = o_P(B_n(\vartheta))$.

Lemma 2.1. *Let $\mathbb{E}[|\varepsilon_k^*|^4] < \infty$ and assume for $\beta^+ > \beta > 0$*

$$|\varphi_\varepsilon(u)|^{-1} = \mathcal{O}\big((1 + |u|)^\beta\big) \quad and \quad |\varphi_\varepsilon'(u)| = \mathcal{O}\big((1 + |u|)^{-\beta - 1}\big)$$

as well as $mb^{2\beta+1} \to \infty$. Then there exists a random variable $\mathcal{E}_b = \mathcal{O}_P\big(1 \vee \frac{1}{m^{1/2}b^{\beta+1}}\big)$ such that for any $s \geqslant 0$ and for any $f \in C^{s+\beta^+}(\mathbb{R})$

$$\left\| \mathcal{F}^{-1}\left[\frac{\varphi_K(bu)}{\varphi_{\varepsilon,m}(u)}\right] * f \right\|_{C^s(\mathbb{R})} = \left\| \mathcal{F}^{-1}\left[\frac{\varphi_K(bu)}{\varphi_{\varepsilon,m}(u)} \mathcal{F} f(u)\right] \right\|_{C^s(\mathbb{R})} \leqslant \mathcal{E}_b \|f\|_{C^{s+\beta^+}(\mathbb{R})}.$$

To prove this lemma, we will show that the linear map $C^{s+\beta^+}(\mathbb{R}) \ni f \mapsto \mathcal{F}^{-1}\big[\varphi_K(bu) \mathcal{F} f(u)/\varphi_{\varepsilon,m}(u)\big] \in C^s(\mathbb{R})$ is bounded. More precisely, the operator norm of the random Fourier multiplier $\varphi_K(bu)/\varphi_{\varepsilon,m}(u)$ can be bounded by the random variable $\mathcal{E}_b$. The condition on the derivative φ_ε' is natural in the context of Fourier multipliers.

Given the assumptions of the lemma the right-hand side of (7) is finite and of course it is continuous in η. Hence, the estimating equation always has a solution for U_n large enough. It does not have to be unique since $\widetilde{f}_b$ is not necessarily non-negative. Nevertheless, any choice converges to the true quantile, assuming the latter is unique. Recalling that we write $\varphi_\varepsilon := \mathcal{F} f_\varepsilon$, the conditions in Lemma 2.1 motivate the definition of the class of error densities f_ε

$$\mathcal{D}^\beta(R, \gamma) := \Big\{ f_\varepsilon \in \mathcal{F}(\infty) \Big| \tfrac{1}{R}(1 + |u|)^{-\beta} \leqslant |\mathcal{F} f_\varepsilon(u)| \leqslant R(1 + |u|)^{-\beta},$$

$$|(\mathcal{F} f_\varepsilon)'(u)| \leqslant R(1 + |u|)^{-1-\beta}, \|x^\gamma f_\varepsilon(x)\|_{L^1} \leqslant R \Big\}$$

for some moment $\gamma > 0$ and we use the same constant R as above for convenience. The upper bound for $|\varphi_\varepsilon(u)|$ in $\mathcal{D}^\beta(R, \gamma)$ is only necessary for the adaptive estimation.

We will need some properties of the kernel to construct our estimator.

Assumption 1. *Let the kernel $K \in L^1(\mathbb{R})$ satisfy*

(i) $\operatorname{supp} \mathcal{F} K \subseteq [-1, 1]$ *and*
(ii) *K has order $\ell \in \mathbb{N}$, i.e., for $k = 1, \ldots \ell$*

$$\int_\mathbb{R} K(x)\,\mathrm{d}x = 1, \quad \int_\mathbb{R} x^k K(x)\,\mathrm{d}x = 0 \quad and \quad \int_\mathbb{R} |K(x)||x|^{\ell+1}\,\mathrm{d}x < \infty.$$

To control the estimation error of $\widetilde{q}_{\tau,b}$, we follow the classical M-estimation approach, or more precisely the Z-estimator approach (cf. van der Vaart (1998)). Let $M(\eta)$ be the deterministic counterpart of $\widetilde{M}_b(\eta)$ defined in (5). The quantities $\widetilde{q}_{\tau,b}$ and q_τ are given by the roots of the derivatives $\widetilde{M}_b'$ and M', respectively. From the Taylor expansion $0 = \widetilde{M}_b'(\widetilde{q}_{\tau,b}) = \widetilde{M}_b'(q_\tau) + (\widetilde{q}_{\tau,b} - q_\tau)\widetilde{M}_b''(q_\tau^*)$ for some intermediate point q_τ^* between q_τ and $\widetilde{q}_{\tau,b}$, we obtain

$$\widetilde{q}_{\tau,b} - q_\tau = -\frac{\int_{-\infty}^{q_\tau}(\widetilde{f}_b(x) - f(x))\,\mathrm{d}x}{2\widetilde{f}_b(q_\tau^*)}. \tag{8}$$

Hence, the error $\widetilde{q}_{\tau,b} - q_\tau$ is affected by the numerator and the denominator in this representation. We now present two propositions that deal with these two terms. These propositions are intrinsic to our analysis, but may also be of interest of their own. The first proposition deals with the numerator in (8) and essentially establishes minimax rates of convergence for distribution deconvolution with unknown error distributions. All convergence rates are determined by the minimum of the sample size n of the observations Y_j and of the sample size m of the observed errors ε_k^*. Therefore, we suppose $n \leqslant m$ throughout. Note that the quotient in (8) might explode if $\widetilde{f}_b(q_\tau^*)$ becomes very small for large stochastic error. Hence, we establish convergence rates as $\mathcal{O}_P$ results.

Proposition 2.2. *Suppose that Assumption 1 holds with $\ell = \langle \alpha \rangle + 1$ and let $b_n^* = n^{-1/(2\alpha + 2(\beta \vee 1/2) + 1)}$. Then for any $\alpha \geqslant 1/2$, $\beta, R, r, \zeta > 0$ and $\gamma \geqslant 4$ we have uniformly over $f \in \mathcal{C}^\alpha(R, r, \zeta)$ and $f_\varepsilon \in \mathcal{D}^\beta(R, \gamma)$ as $n \to \infty$,*

$$\Big| \int_{-\infty}^{q_\tau} (\widetilde{f}_{b_n^*}(x) - f(x)) \, \mathrm{d}x \Big| = \mathcal{O}_P(\psi_n(\alpha, \beta)),$$

where

$$\psi_n(\alpha, \beta) := \begin{cases} n^{-1/2}, & \text{for } \beta \in (0, 1/2), \\ (\log n / n)^{1/2}, & \text{for } \beta = 1/2, \\ n^{-(\alpha+1)/(2\alpha + 2\beta + 1)}, & \text{for } \beta > 1/2. \end{cases} \tag{9}$$

In view of the lower bounds stated in Fan (1991), as long as $n \leq m$, the rates above are optimal and estimating the distribution function by integrating a density deconvolution estimator is a minimax optimal procedure. Hence, this proposition closes the gap reported in Fan (1991) and further extends the results to the case of unknown error distributions. In the studies of density estimation with unknown error distribution, for instance Neumann (1997); Johannes (2009), the sample size m of the error (ε_k^*) can be of smaller order than the number n of the observations (Y_j) to obtain optimal rates. This is because the risk of estimating of φ_ε profits from the decay of the characteristic function of X_j. Assuming local regularity on f only, its Fourier transform will not decay fast in general such that this effect does not occur. Consequently, it is natural that the minimax rates are indeed determined by $n \wedge m$ (cf. the analysis of (27) below).

Next we would like to understand better the denominator of (8). Lounici and Nickl (2011) proved uniform risk bounds for the deconvolution wavelet estimator on the whole real line. However, on a bounded interval, which is sufficient for our purpose, uniform convergence of the deconvolution estimator $\widetilde{f}_b$ can be proved more elementarily. Note that with $b_n = (\log n / n)^{1/(2\alpha + 2\beta + 1)}$ the following proposition yields the minimax rate $(\log n / n)^{\alpha/(2\alpha + 2\beta + 1)}$ of the L^∞-loss.

Proposition 2.3. *Grant Assumption 1 with $\ell = \langle \alpha \rangle$. For any $\alpha, \beta, R, r, \zeta > 0$ and $\gamma \geqslant 0$ we have uniformly over $f \in \mathcal{C}^\alpha(R, r, \zeta)$ and $f_\varepsilon \in \mathcal{D}^\beta(R, \gamma)$ as $n \to \infty$,*

$$\sup_{x \in (-\zeta, \zeta)} |\widetilde{f}_b(x + q_\tau) - f(x + q_\tau)| = \mathcal{O}_P\Big(b^\alpha + \Big(\frac{\log n}{nb^{2\beta+1}}\Big)^{1/2}\Big).$$

In particular, if $b_n \to 0$ and $nb_n^{2\beta+1}/\log n \to \infty$ as $n \to \infty$, $\widetilde{f}_{b_n}$ is a uniformly consistent estimator.

The density deconvolution estimator is then locally uniformly consistent. The two propositions above are the building blocks for the first main result of this paper established in the following theorem.

Theorem 2.4. *Suppose that Assumption 1 holds with $\ell = \langle \alpha \rangle + 1$. Let $\widetilde{q}_{\tau,b}$ be the quantile estimator defined in (5) associated with $b_n^* = n^{-1/(2\alpha + 2(\beta \vee 1/2) + 1)}$. Then for any $\alpha \geqslant 1/2$, $\beta, R, r, \zeta > 0$ and $\gamma \geqslant 4$ we have uniformly over $f \in \mathcal{C}^\alpha(R, r, \zeta)$ and $f_\varepsilon \in \mathcal{D}^\beta(R, \gamma)$ as $n \to \infty$,*

$$|\widetilde{q}_{\tau,b_n^*} - q_\tau| = \mathcal{O}_P(\psi_n(\alpha, \beta))$$

where $\psi_n(\alpha, \beta)$ is given in (9).

Remarks 2.5.

(i) Although we do not provide the lower bounds for quantile estimation, it is clear that the above rate is the optimal. It is the same rate as the one achieved in Proposition 2.2 which deals with the estimation of the distribution function. A better rate cannot be expected (cf. the case of no measurement errors).

(ii) In order to prove the theorem, we apply a smooth truncation function a_s to decompose the error into

$$\int_{-\infty}^{q_\tau} (\widetilde{f}_b(x) - f(x)) \, \mathrm{d}x$$

$$= \underbrace{\int_{-\infty}^{q_\tau} (K_b * f(x) - f(x)) \, \mathrm{d}x}_{\text{deterministic error}} + \underbrace{\int_{-\infty}^{q_\tau} a_s(x + q_\tau)(\widetilde{f}_b(x) - K_b * f(x)) \, \mathrm{d}x}_{\text{singular part of stochastic error}}$$

$$+ \underbrace{\int_{-\infty}^{q_\tau} (1 - a_s)(x + q_\tau)(\widetilde{f}_b(x) - K_b * f(x)) \, \mathrm{d}x}_{\text{continuous part of stochastic error}} \tag{10}$$

with the usual notation $K_b(\cdot) = b^{-1}K(\cdot/b)$. The function a_s can be chosen such that it has compact support and satisfies $(\mathbb{1}_{(-\infty,0]} - a_s) \in C^\infty(\mathbb{R})$. Similar to the classical bias-variance trade-off, the deterministic error and singular part of the stochastic error will determine the rate. The continuous part, however, corresponds to the estimation error of a smooth (but not integrable) functional of the density. If the error distribution would be known, it would be of order $n^{-1/2}$. For unknown errors we use Lemma 2.1.

(iii) Dealing with unknown error distributions leads to the requirement of $\gamma \geqslant 4$, i.e., that f_ε possesses at least 4 moments. In view of the analysis by Neumann and Reiß (2009) this assumption also implies uniform convergence of $\varphi_{\varepsilon,m}$. As seen in Lemma 2.1, our estimate of the operator norm of the random Fourier multiplier $\mathcal{F}^{-1}[\varphi_K(bu)/\varphi_{\varepsilon,m}(u)]$ is of order $\mathcal{O}_P((nb^{\beta+1})^{-1})$. This might be larger than the operator norm of the unknown deconvolution operator $\mathcal{F}^{-1}[1/\varphi_\varepsilon(u)]$ which is uniformly bounded. Yet, for $\alpha \geqslant 1/2$ the additional error that appears in the continuous part of stochastic error in (10) is negligible. More generally, the continuous part is of smaller or of the same order as the singular part whenever $n^{-1}b_n^{-2\beta \wedge 1 - 2}$ is bounded.

We finish this section by providing the minimax rates for estimating the distribution function and the quantiles for the case of known error distributions. As above the estimators are given by plug-in, using the classical density estimator

$$\widehat{f}_b(x) := \mathcal{F}^{-1}\left[\frac{\varphi_n(u)\varphi_K(bu)}{\varphi_\varepsilon(u)}\right](x), \quad x \in \mathbb{R}. \tag{11}$$

Due to the known φ_ε the mathematical analysis is simpler and thus we need weaker assumptions.

Corollary 2.6. *Suppose that the error distribution is known. Let Assumption 1 hold with $\ell = \langle \alpha \rangle + 1$. Let $\widehat{q}_{\tau,b}$ be the quantile estimator based on the density deconvolution estimator (11) associated with $b_n^* = n^{-1/(2\alpha+2(\beta \vee 1/2)+1)}$. Then for any $\alpha, \beta, R, r, \zeta > 0$ and $\gamma \geqslant 0$ we obtain uniformly over $f \in \mathcal{C}^\alpha(R,r,\zeta)$ and $f_\varepsilon \in \mathcal{D}^\beta(R,\gamma)$ as $n \to \infty$,*

$$\left| \int_{-\infty}^{q_\tau} (\widehat{f}_{b_n^*}(x) - f(x)) \, \mathrm{d}x \right| = \mathcal{O}_P(\psi_n(\alpha,\beta)),$$

$$|\widehat{q}_{\tau,b_n^*} - q_\tau| = \mathcal{O}_P(\psi_n(\alpha,\beta)),$$

where $\psi_n(\alpha,\beta)$ is given (9).

Remarks 2.7.

(i) We want to stress again that the only global condition on the density f is uniform boundedness. A tail condition similar to the one of Fan (1991) is not necessary. In contrast to Dattner, Goldenshluger and Juditsky (2011), the smoothness is measured locally in a Hölder scale and not globally by decay conditions of the Fourier transform of f. The former is more natural since both, the distribution function and the quantile function are estimated pointwise.

(ii) Using the known φ_ε, we do not need to estimate the deconvolution operator and thus there is no additional error. Consequently, we do not need a moment assumption on the error distribution and the continuous part of the stochastic error in (10) is of order $n^{-1/2}$ implying that the minimax convergence rates hold true for all $\alpha > 0$.

3. Adaptive estimation

The choice of the bandwidth b is crucial in applications. Therefore, we develop a fully data-driven procedure to determine a reasonable bandwidth. We follow the approach which was originally stated by Lepskiĭ (1990). More precisely, we use the version proposed in Goldenshluger and Nemirovski (1997).

To this end, we consider the family of estimators $\{\widetilde{q}_{\tau,b}, b \in \mathcal{B}_n\}$ where $\widetilde{q}_{\tau,b}$ is defined in (5) and $\mathcal{B}_n$ is a finite set of bandwidths. Typically, the bandwidths are chosen geometrically growing. In view of the error representation (8) it is important that $\widetilde{f}_b(\widetilde{q}_{\tau,b})$ is a consistent estimator of $f(q_\tau)$ for all $b \in \mathcal{B}_n$. Therefore, conditions on the bandwidth as in Proposition 2.3 are necessary for the whole set $\mathcal{B}_n$. Moreover, we keep to the assumption $\alpha > 1/2$ such that the additional error due to bounding the random Fourier multiplier is negligible. We will prove convergence rates for the adaptive procedure under the following

Assumption 2. *Let the set $\mathcal{B}_n := \{b_{n,j}, j = 1, ..., N_n\}$ consists of a monotone increasing sequence of bandwidths such that $b_{n,j+1}/b_{n,j}$ is uniformly bounded in $j = 1, \ldots, N_n$ and $n \geqslant 1$. For $n \to \infty$ suppose*

$$N_n \lesssim \log n, \qquad (\log n)^2 b_{n,N_n} \to 0 \qquad and \qquad n b_{n,1}^{2\beta+2} \to \infty.$$

Moreover, the optimal bandwidth $b_n^* = n^{-1/(2\alpha+2(\beta\vee 1/2)+1)}$ has to be contained in the interval $[b_{n,1}, b_{n,N_n}]$.

Obviously, Assumption 2 depends on the true but unknown degree of ill-posedness β. Note that in our case the lower bound for the bandwidth is not determined by the variance of the quantile estimator itself but by the variance of the density estimator and the minimal smoothing which results from $\alpha > 1/2$. Inspired by Comte and Lacour (2011), we propose the following construction of a feasible set $\mathcal{B}_n$.

Lemma 3.1. *Defining for $\Lambda_n := \{1, 1/\sqrt{2}, \ldots, 1/\sqrt{n}\}$*

$$\widetilde{b}_{n,min} := \min\left\{ b \in \Lambda_n : \frac{1}{2} \leqslant \left(\frac{\log n}{n}\right)^{1/2} \int_{-1/b}^{1/b} |\varphi_{\varepsilon,m}(u)|^{-1} \, \mathrm{d}u \leqslant 1 \right\} \tag{12}$$

and $L_n := (\widetilde{b}_{n,min}(\log n)^3)^{-1/\log n}$, let $\mathcal{B}_n$ be given by

$$N_n = \lfloor \log n \rfloor \quad and \quad b_{n,1} = \widetilde{b}_{n,min}, \quad b_{n,j} = L_n^j b_{n,1}, \quad j = 2, \ldots, N_n.$$

Then, $\mathcal{B}_n$ satisfies Assumption 2 for all $f \in \mathcal{C}^\alpha(R, r, \zeta)$ and $f_\varepsilon \in \mathcal{D}^\beta(R, \gamma)$ with $\alpha > 1/2$ and $\beta > 0$

Given the bandwidth set, the adaptive estimator is obtained by selection from the family of estimators $\{\widetilde{q}_{\tau,b}, b \in \mathcal{B}_n\}$. As proposed by Lepskiĭ (1990) the adaptive choice should mimic the trade-off between deterministic error and stochastic error. The adaptive choice will be given by the largest bandwidth such that the intersection of all confidence sets, which corresponds to smaller bandwidths, is non-empty. As discussed above it is sufficient to consider the singular part of the stochastic error in (10) only. To estimate the variance of $\widetilde{q}_{\tau,b}$ corresponding to the latter, we define for some $\delta > 0$

$$\widetilde{\Sigma}_b := \frac{(\sqrt{2}+\delta)\sqrt{\log\log n}\max_{\mu \geqslant b}\widetilde{\sigma}_{\mu,X} + (\delta\log n)^3\max_{\mu \geqslant b}\widetilde{\sigma}_{\mu,\varepsilon}}{|\widetilde{f}_b(\widetilde{q}_{\tau,b})|}, \tag{13}$$

with the truncation function a_s from decomposition (10) and

$$\widetilde{\sigma}_{b,X}^2 = \frac{1}{n^2}\sum_{j=1}^{n}\left(\int_{-\infty}^{0} a_s(x)\mathcal{F}^{-1}\left[\frac{\varphi_K(bu)e^{iuY_j}}{\varphi_{\varepsilon,m}(u)}\right](x+\widetilde{q}_{\tau,b})\,\mathrm{d}x\right)^2 \quad and$$

$$\widetilde{\sigma}_{b,\varepsilon}^2 = \frac{1}{4\pi^2 m}\int_{-1/b}^{1/b}|\varphi_K(bu)|\left|\frac{\varphi_n(u)}{\varphi_{\varepsilon,m}(u)}\right|^2 \mathrm{d}u \int_{-1/b}^{1/b}|\varphi_K(bu)|\left|\frac{\mathcal{F}a_s(u)}{\varphi_{\varepsilon,m}(u)}\right|^2 \mathrm{d}u.$$

Note that we apply a monotonization in the numerator of $\widetilde{\Sigma}_b$ by taking maxima of $\widetilde{\sigma}_{\mu,X}$ and $\widetilde{\sigma}_{\mu,\varepsilon}$. With $\widetilde{\Sigma}_b$ at hand the adaptive estimator is given according to the following rule. Define for any $b \in \mathcal{B}_n$

$$\mathcal{U}_b := [\widetilde{q}_{\tau,b} - \widetilde{\Sigma}_b, \widetilde{q}_{\tau,b} + \widetilde{\Sigma}_b]. \tag{14}$$

The adaptive estimator is given by

$$\widetilde{q}_\tau := \widetilde{q}_{\tau,\widetilde{b}_n^*} \quad \text{with} \quad \widetilde{b}_n^* := \max\left\{ b \in \mathcal{B}_n \middle| \bigcap_{\mu \leq b, \mu \in \mathcal{B}_n} \mathcal{U}_\mu \neq \emptyset \right\}. \tag{15}$$

Note that $\widetilde{b}_n^*$ is well-defined since the intersection in (15) is non-empty for $b = b_{n,1}$. The following theorem shows that this estimator achieves the minimax rate up to a logarithmic factor. As usual the proof relies on a comparison with an oracle-type choice of the bandwidth. However, all ingredients have to be estimated and without assuming independence of Y_j and ε_k^*, which requires special attention.

Theorem 3.2. *Grant Assumptions 1 and 2 with $\ell \geqslant \langle \alpha \rangle + 1$. Then for any $\alpha > 1/2$, $\beta, R, r, \zeta > 0$ and $\gamma \geqslant 4$ the estimator $\widetilde{q}_\tau$ as defined in (15) satisfies uniformly over $f \in \mathcal{C}^\alpha(R, r, \zeta)$ and $f_\varepsilon \in \mathcal{D}^\beta(R, \gamma)$ as $n \to \infty$,*

$$|\widetilde{q}_\tau - q_\tau| = \mathcal{O}_P\left(\left(\sqrt{\log \log n} + \left(\log n^\delta\right)^3\right)\psi_n(\alpha, \beta)\right),$$

where $\psi_n(\alpha, \beta)$ is given in (9).

Remarks 3.3.

(i) Since Y_j and ε_k^* are not independent, we have to bound the stochastic error of $\widetilde{q}_{\tau,b}$ with use of the Cauchy–Schwarz inequality to separate the error terms. To estimate the remaining factors, we take into account the term $(\log n^\delta)^3$. If the error density were known, this term would not appear. The $\sqrt{\log \log n}$ is the additional loss for $\mathcal{O}_P$-adaptivity which seems unavoidable.

(ii) To estimate the distribution function, a very similar procedure can be employed because the bias-variance trade-off is determined by the estimation of the distribution function anyway. The only difference will be to replace the denominator of $\widetilde{\Sigma}_b$ by $1/2$ which results in

$$(2\sqrt{2} + \delta)\sqrt{\log \log n} \max_{\mu \geqslant b} \widetilde{\sigma}_{\mu,X} + (\delta \log n)^3 \max_{\mu \geqslant b} \widetilde{\sigma}_{\mu,\varepsilon},$$

for some $\delta > 0$. The proof for an analogous result as Theorem 3.2 follows easily.

4. Numerical results

4.1. Simulation study

We now illustrate the implementation of the adaptive estimation procedure of Sections 3. Our small simulation study serves as a proof of viability of our proposed method.

We run 500 Monte Carlo simulations for four experimental setups. The sample size is set to $n = 1000$ and the external sample of the directly observed error is set to $m = 1000$ as well (here the external sample is independent of the main one). We consider $\Gamma(1,1)$ and $\Gamma(2,1)$ for the distribution of X where $\Gamma(k, \eta)$ denotes the gamma distribution with shape parameter k and scale η such that its density is $(\Gamma(\alpha)\eta^k)^{-1} x^{k-1} \exp\{-x/\eta\} \mathbb{1}_{\{x \geqslant 0\}}$. Note that the shape k of the gamma distribution determines the global smoothness of the density while our convergence rates depend on local smoothness α which may be larger. For the error distribution we consider standard Laplace distribution ($\beta = 2$) and the convolution of a standard Laplace with itself ($\beta = 4$).

	$k = 1, \beta = 2$	$k = 2, \beta = 2$	$k = 1, \beta = 4$	$k = 2, \beta = 4$
$\tau = 0.25$				
RMSE	0.279 (0.296)	0.146 (0.239)	0.295 (0.651)	0.169 (0.516)
$\tau = 0.5$				
RMSE	0.233 (0.170)	0.187 (0.143)	0.246 (0.239)	0.202 (0.204)
$\tau = 0.75$				
RMSE	0.453 (0.541)	0.245 (0.436)	0.470 (0.923)	0.256 (0.761)

The target quantiles of interest are q_τ with $\tau = 0.25, 0.5, 0.75$. A kernel with flat-top Fourier transform was chosen for the estimator (see e.g., McMurry and Politis (2004)) and the adaptive algorithm was implemented on a geometrically growing grid. Usually, applying this adaptive scheme requires an additional tuning of the algorithm (see e.g. Spokoiny and Vial (2009)). However, for the specific setup considered here the adaptive procedure seems to be robust and was implemented exactly as defined in (14)-(15). In the real data example in the next subsection we compare the adaptive estimator to the "naive" quantile estimator given by the inverse of the empirical distribution function of the observations Y. Thus we also applied the naive estimator in our simulations. The results of this simulation study are given in Table 1. We can see that the results support the theory - the empirical root mean squared error (RMSE) is higher for $\beta = 4$ than for $\beta = 2$. Also, we can see that the RMSE is lower for $k = 2$ than for $k = 1$ since the gamma distribution with larger shape parameter is smoother in our context. For median estimation the naive seems to perform quiet well in this experimental setup.

4.2. Real data example

High blood pressure is a direct cause of serious cardiovascular disease (Kannel et al. (1995)) and determining reference values for physicians is important. In particular, estimating percentiles of systolic and diastolic blood pressure by sex, race or ethnicity, age, etc. is of substantial interest. However, blood pressure is known to be measured with additional error which needs to be addressed in its analysis (see e.g., Frese, Fick and Sadowsky (2011)). Therefore, measurement errors should be taken into account, otherwise quantile estimates based on the observed blood pressure measurements would be biased.

The resulting adaptive distribution and quantiles estimates are displayed in Figure 2. We illustrate our method using data from the Framingham Heart Study (Carroll et al. (2006)). This study consists of a series of exams taken two years apart where systolic blood pressure (SBP) measurements of $1,615$ men aged $31 - 65$ were taken. These data were used as an illustration for density deconvolution by Stirnemann, Comte and Samson (2012) and for distribution deconvolution by Dattner and Reiser (2013). We denote by $Y_{j,1}$ and $Y_{j,2}$ the two repeated measures of SBP for each individual j at two different exams and denote by X_j the long-term average SBP of individual j. Then we assume that

$$Y_{j,1} = X_j + \varepsilon_{j,1}, \qquad Y_{j,2} = X_j + \varepsilon_{j,2},$$

for individuals $j = 1, ..., n$. Following Carroll et al. (2006), we use the average of the two exams $Y_j' = (Y_{j,1} + Y_{j,2})/2$, so that the model in our case is

$$Y_j' = X_j + \varepsilon_j', \tag{16}$$

where $\varepsilon_j' = (\varepsilon_{j,1} + \varepsilon_{j,2})/2$.

Taking advantage of the repeated measures setup, we can avoid parametric assumptions regarding the distribution of the errors. The only assumption we will make is that the distribution of the measurement error is symmetric around zero and does not vanish. We then set $\varepsilon_j^* = (Y_{j,1} - Y_{j,2})/2$
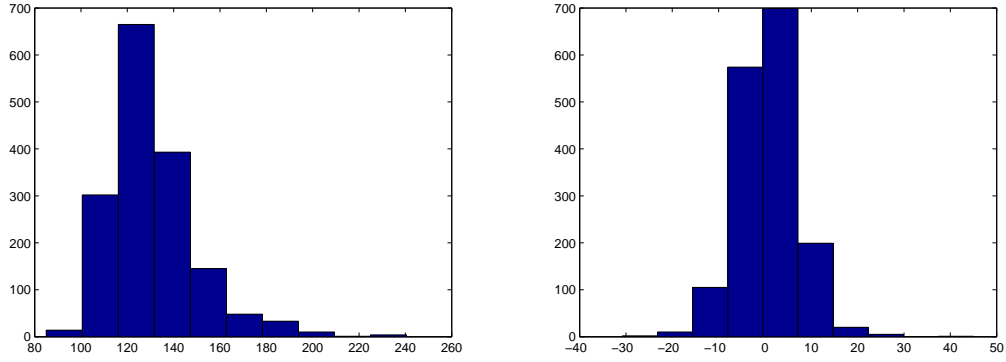
FIG 1. *Average systolic blood pressure Y' (left) and the errors ε^* (right) over the two measurements from the two visits of $1,615$ men aged $31 - 65$ from the Framingham Heart Study.*
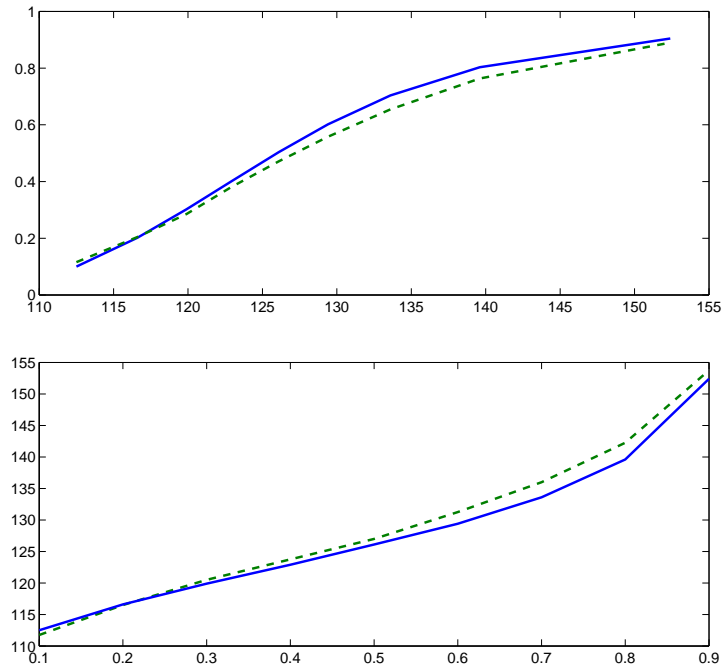

FIG 2. *Distribution function estimation (top) and quantiles estimates (bottom) for systolic blood pressure of $1,615$ men aged $31 - 65$ from the Framingham Heart Study. Solid line for the adaptive estimator and dashed line for the naive estimator.*

and note that under the symmetry assumption it is distributed as ε'_j. We emphasize the fact that our theoretical results do not require that the sample ε^*_j will be independent from that of the Y'_j. Thus, we estimate the characteristic function of the error ε' by

$$\varphi_{\varepsilon',m}(u) = \frac{1}{m} \sum_{j=1}^{m} \exp(iu\varepsilon^*_j).$$

Histograms of Y' and ε^* are presented in Figure 1.

With the adaptive chosen bandwidth, the Fourier transform of the kernel function K_b is supported on $[-0.1, 0.1]$ for both distribution and quantiles estimation. It seems that there is no much difference between the naive and adaptive estimates for smaller quantiles. However, the observed differences are larger for larger quantiles. Although further analysis of the data may and should be done for inference purposes, we do not pursue them here since our goal is merely to show the applicability of the adaptive procedure to real data.

Appendix A: Proofs for Section 2

In the sequel we will use the deterministic Landau symbols $\mathcal{O}$ and o. For convenience we will write $A_n(\vartheta) \lesssim B_n(\vartheta)$ if $A_n(\vartheta) = \mathcal{O}(B_n(\vartheta))$ as well. For a better readability we throughout assume $\beta \neq 1/2$. In the special case $\beta = 1/2$ the order of the stochastic error will be $(\log n/n)^{1/2}$ which can be easily seen below in the bounds (28) and (30). For the sake of clarity of our arguments we distinguish between n and m in this section. However, all rates are governed by $n \wedge m$. The subscript n at the bandwidth will be omitted.

Since $1/\varphi_{\varepsilon,m}$ might explode for large stochastic errors we need the following lemma. The moment assumption on f_ε corresponds to the condition $\gamma > 2$ on the set $\mathcal{D}^\beta(R, \gamma)$.

Lemma A.1. *Suppose $\mathbb{E}[|\varepsilon^*_k|^{2+\delta}] < \infty$ for some $\delta > 0$. Let $T_m \to \infty$ be an increasing sequence satisfying $m^{1/2} \inf_{u \in [-T_m, T_m]} |\varphi_\varepsilon(u)| \gtrsim (\log T_m)^2$, then for any $p < 2$*

$$P\Big(\inf_{u \in [-T_m, T_m]} |\varphi_{\varepsilon,m}(u)| < m^{-1/2}(\log T_m)^p\Big) = o(1) \quad as \quad m \to \infty.$$

Proof. The triangle inequality, the assumption on T_m and Markov's inequality yield for some constant $D > 0$ and for m as well as T_m large enough

$$P\Big(\inf_{u \in [-T_m, T_m]} |\varphi_{\varepsilon,m}(u)| < m^{-1/2}(\log T_m)^p\Big)$$

$$\leqslant P\Big(\exists u \in [-T_m, T_m] : |\varphi_\varepsilon(u) - \varphi_{\varepsilon,m}(u)| > |\varphi_\varepsilon(u)| - m^{-1/2}(\log T_m)^p\Big)$$

$$\leqslant P\Big(\sup_{u \in [-T_m, T_m]} |\varphi_\varepsilon(u) - \varphi_{\varepsilon,m}(u)| > \inf_{u \in [-T_m, T_m]} |\varphi_\varepsilon(u)| - m^{-1/2}(\log T_m)^p\Big)$$

$$\leqslant P\Big(\sup_{u \in [-T_m, T_m]} m^{1/2}|\varphi_\varepsilon(u) - \varphi_{\varepsilon,m}(u)| > D(\log T_m)^2 - (\log T_m)^p\Big)$$

$$\leqslant \frac{2}{D(\log T_m)^2} \mathbb{E}\Big[\sup_{u \in [-T_m, T_m]} m^{1/2}|\varphi_\varepsilon(u) - \varphi_{\varepsilon,m}(u)|\Big].$$

Noting $\mathbb{1}_{[-T_m, T_m]}(u) \leqslant w(u)/w(T_m)$ for $w(u) := (\log(e + |u|))^{-1/2-\eta}$ for some $\eta \in (0, 1/2)$, the above display can be bounded by

$$\frac{2}{Dw(T_m)(\log T_m)^2} \mathbb{E}\Big[\sup_{u \in \mathbb{R}} m^{1/2}w(u)\big|\varphi_\varepsilon(u) - \varphi_{\varepsilon,m}(u)\big|\Big] \lesssim (\log T_m)^{-3/2+\eta} \tag{17}$$

where the expectation is bounded by applying Theorem 4.1 in Neumann and Reiß (2009). □

To ensure consistency of the density estimator, we have to assume frequently $(n \wedge m)b^{2\beta+1} \to \infty$. Since this implies $m^{1/2} \inf_{u \in [-1/b,1/b]} |\varphi_\varepsilon(u)| \gtrsim |\log b|^2$ for $f \in \mathcal{D}^\beta(R, \gamma)$ and any bandwidth which increases polynomially in n, Lemma A.1 can be applied to $T_m = 1/b$. Hence, under this assumption the probability of the event

$$B_\varepsilon(b) := \left\{ \inf_{u \in [-1/b,1/b]} |\varphi_{\varepsilon,m}(u)| \geqslant m^{-1/2}|\log b|^{3/2} \right\} \tag{18}$$

tends to one. Therefore, it is sufficient to control terms on $B_\varepsilon(b)$ only. Frequently, the weaker estimate $|\varphi_{\varepsilon,m}(u)| \geqslant m^{-1/2}$ for $|u| \leqslant 1/b$ will be enough.

A.1. Proof of Lemma 2.1

Note that the assumptions on φ_ε imply $|(\varphi_\varepsilon^{-1})'(u)| \lesssim (1+|u|)^{\beta-1}$ as well as $|\varphi_\varepsilon^{-1}(u)| \lesssim (1+|u|)^\beta, u \in \mathbb{R}$. On these assumptions Söhl and Trabs (2012) have shown that $(1+iu)^{-\beta}/\varphi_\varepsilon(u)$ is a Fourier multiplier on Besov spaces. Due to the regularization with the kernel, $\varphi_K(bu)/\varphi_{\varepsilon,m}(u)$ behaves basically like $1/\varphi_\varepsilon$ and thus we can derive mapping properties of the random Fourier multiplier

$$\psi(u) := (1+iu)^{-\beta}\frac{\varphi_K(bu)}{\varphi_{\varepsilon,m}(u)}, \quad u \in \mathbb{R}.$$

On $B_\varepsilon(b)$, as defined in (18), we will check Hörmander type conditions and derive an upper bound for the operator norm of $\psi(u)$. More precisely, we apply Corollary 4.13 by Girardi and Weis (2003) with $p = 2$, $l = 1$. Hence, we have to determine a suitable constant $A_\psi > 0$ satisfying

$$\max_{l \in \{0,1\}} \left(\int_{[-2,2]} |\psi^{(l)}(u)|^2 \, \mathrm{d}u \right)^{1/2} \leqslant A_\psi \quad \text{and}$$

$$\max_{l \in \{0,1\}} \sup_{T \in [0,\infty)} T^{l-1/2}\left(\int_{T \leqslant |u| \leqslant 4T} |\psi^{(l)}(u)|^2 \, \mathrm{d}u \right)^{1/2} \leqslant A_\psi. \tag{19}$$

To find A_ψ, we note that

$$\frac{1}{|\varphi_{\varepsilon,m}(u)|^p} \leqslant \frac{p}{|\varphi_\varepsilon(u)|^p} + \frac{p|\varphi_{\varepsilon,m}(u) - \varphi_\varepsilon(u)|^p}{|\varphi_\varepsilon(u)\varphi_{\varepsilon,m}(u)|^p}, \quad \text{for } p \in \{1, 2\} \tag{20}$$

and thus on $B_\varepsilon(b)$

$$\frac{1}{|\varphi_{\varepsilon,m}(u)|} \leqslant \frac{1 + \Delta_m(u)}{|\varphi_\varepsilon(u)|}, \qquad \Delta_m(u) := \frac{m^{1/2}}{|\log b|^{3/2}}|\varphi_{\varepsilon,m}(u) - \varphi_\varepsilon(u)|,$$

By $f_\varepsilon \in \mathcal{D}^\beta(R, \gamma)$ we conclude

$$|\psi(u)| \leqslant \frac{|\varphi_K(bu)|(1 + \Delta_m(u))}{(1+u^2)^{\beta/2}|\varphi_\varepsilon(u)|} \lesssim (1 + \Delta_m(u))\mathbb{1}_{[-1/b,1/b]}(u). \tag{21}$$

Concerning the derivative, we estimate $b \leqslant 2(1+|u|)^{-1}$ for $|u| \leqslant 1/b$ and $b < 1/2$ and consequently by $|\varphi'_\varepsilon(u)/\varphi_\varepsilon(u)| \lesssim (1+|u|)^{-1}$

$$|\psi'(u)| \leqslant (\beta+1)(1+u^2)^{-(\beta+1)/2}\left|\frac{\varphi_K(bu)}{\varphi_{\varepsilon,m}(u)}\right| + b(1+u^2)^{-\beta/2}\left|\frac{\varphi'_K(bu)}{\varphi_{\varepsilon,m}(u)}\right|$$

$$+ (1+u^2)^{-\beta/2}\left|\frac{\varphi'_{\varepsilon,m}(u)}{\varphi_{\varepsilon,m}(u)}\frac{\varphi_K(bu)}{\varphi_{\varepsilon,m}(u)}\right|$$

$$\lesssim \frac{|\psi(u)|}{1+|u|} + |\psi(u)|\left|\frac{\varphi'_{\varepsilon,m}(u)}{\varphi_{\varepsilon,m}(u)}\right|$$

$$\lesssim \frac{1 + \Delta_m(u)}{1 + |u|} \Big(1 + \Big| \frac{\varphi'_{\varepsilon,m}(u)}{\varphi'_\varepsilon(u)} \frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)} \Big| \Big) \mathbb{1}_{[-1/b,1/b]}(u)$$

$$\lesssim \frac{(1 + \Delta_m(u))^2}{1 + |u|} \Big(1 + \frac{|\varphi'_{\varepsilon,m}(u) - \varphi'_\varepsilon(u)|}{|\varphi'_\varepsilon(u)|} \Big) \mathbb{1}_{[-1/b,1/b]}(u)$$

$$\lesssim \frac{1 + \Delta_m^2(u)}{1 + |u|} \Big(1 + (1 + |u|)^{\beta+1} |\varphi'_{\varepsilon,m}(u) - \varphi'_\varepsilon(u)| \Big) \mathbb{1}_{[-1/b,1/b]}(u). \tag{22}$$

With these bounds at hand we can show now (19). For $l = 0$ the estimate (21) and $1/T \lesssim (1+|u|)^{-1}$ for $|u| \leqslant 4T$ yield

$$\int_{-2}^{2} |\psi(u)|^2 \, du \lesssim \int_{-2}^{2} (1 + \Delta_m^2(u)) \mathbb{1}_{[-1/b,1/b]}(u) \, du,$$

$$\frac{1}{T} \int_{T \leqslant |u| \leqslant 4T} |\psi(u)|^2 \, du \lesssim \frac{1}{T} \int_{T \leqslant |u| \leqslant 4T} (1 + \Delta_m^2(u)) \mathbb{1}_{[-1/b,1/b]}(u) \, du$$

$$\lesssim 1 + \int_{-1/b}^{1/b} (1 + |u|)^{-1} \Delta_m^2(u) \, du.$$

Hence, the conditions (19) for $l = 0$ are satisfied for A_ψ of the order

$$\Big(1 + \int_{-1/b}^{1/b} (1 + |u|)^{-1} \Delta_m^2(u) \, du \Big)^{1/2}.$$

For $l = 1$ we verify by (22) and $T \leqslant (1 + |u|)$ for $|u| > T$

$$\int_{-2}^{2} |\psi'(u)|^2 \, du \lesssim \int_{-2}^{2} \big(1 + \Delta_m^4(u) \big) \big(1 + (1 + |u|)^{2\beta+2} |\varphi'_{\varepsilon,m}(u) - \varphi'_\varepsilon(u)|^2 \big) \, du \quad \text{and}$$

$$T \int_{T \leqslant |u| \leqslant 4T} |\psi'(u)|^2 \, du$$

$$\lesssim \int_{T \leqslant |u| \leqslant 4T} \frac{T \, du}{(1 + |u|)^2} + \int_{-1/b}^{1/b} \Big(\frac{\Delta_m^4(u)}{1 + |u|} + (1 + \Delta_m^4(u))(1 + |u|)^{2\beta+1} |\varphi'_{\varepsilon,m}(u) - \varphi'_\varepsilon(u)|^2 \Big) \, du$$

$$\lesssim 1 + \int_{-1/b}^{1/b} \Big(\frac{\Delta_m^4(u)}{1 + |u|} + (1 + \Delta_m^4(u))(1 + |u|)^{2\beta+1} |\varphi'_{\varepsilon,m}(u) - \varphi'_\varepsilon(u)|^2 \Big) \, du.$$

Therefore, we find a constant $A' > 0$, depending only on R, β, such that (19) holds for

$$A_\psi := A' \Big(1 + \int_{-1/b}^{1/b} \Big(\frac{\Delta_m^2(u) + \Delta_m^4(u)}{1 + |u|}$$

$$+ (1 + \Delta_m^4(u))(1 + |u|)^{2\beta+1} |\varphi'_{\varepsilon,m}(u) - \varphi'_\varepsilon(u)|^2 \Big) \, du \Big)^{1/2}. \tag{23}$$

Now, the conditions (19) imply that ψ is indeed a Fourier multiplier on $B_\varepsilon(b)$ and thus by Theorem 4.8 and Corollary 4.13 by Girardi and Weis (2003) there is a universal constant $C > 0$ such that for all $\eta > 0$ and $f \in C^{s+\beta+\eta}(\mathbb{R})$

$$\Big\| \mathcal{F}^{-1} \Big[\frac{\varphi_K(bu)}{\varphi_{\varepsilon,m}(u)} \Big] * f \Big\|_{C^s} = \Big\| \mathcal{F}^{-1} \Big[\frac{\varphi_K(bu)}{\varphi_{\varepsilon,m}(u)} \mathcal{F} f \Big] \Big\|_{C^s}$$

$$\leqslant C A_\psi \Big\| \mathcal{F}^{-1} \big[(1 + iu)^\beta \mathcal{F} f \big] \Big\|_{C^{s+\eta}}.$$

Since the Fourier multiplier $(1 + iu)^\beta$ is an isomorphism from $C^{s+\beta}(\mathbb{R})$ onto $C^s(\mathbb{R})$ (Triebel, 2010, Thm. 2.3.8), there is another universal constant $C' > 0$ such that the first assertion of the lemma follows:

$$\Big\| \mathcal{F}^{-1} \Big[\frac{\varphi_K(bu)}{\varphi_{\varepsilon,m}(u)} \mathcal{F} f \Big] \Big\|_{C^s} \leqslant \mathcal{E}_b \|f\|_{C^{s+\beta+\eta}} \quad \text{with} \quad \mathcal{E}_b := C' A_\psi.$$

To bound $\mathcal{E}_b$, we apply Markov's inequality on A_ψ from (23). The inequality by Rosenthal (1970) yields

$$\mathbb{E}\left[m^{p/2}\left|\varphi_{\varepsilon,m}^{(l)}(u) - \varphi_\varepsilon^{(l)}(u)\right|^p\right] < \infty$$

for $l = 0$ and $p \in \mathbb{N}$ as well as $l = 1$ and $p \in \{1, \ldots, 4\}$. Combined with the Cauchy–Schwarz inequality, we obtain

$$
\begin{aligned}
P\Big(&B_\varepsilon(b) \cap \Big\{\mathcal{E}_b > \frac{c^{1/2}}{m^{1/2}b^{\beta+1} \wedge 1}\Big\}\Big) \\
&= P\Big(B_\varepsilon(b) \cap \Big\{A_\psi > \frac{c^{1/2}}{C'(m^{1/2}b^{\beta+1} \wedge 1)}\Big\}\Big) \\
&\leqslant c^{-1}C'^2(mb^{2\beta+2} \wedge 1)\,\mathbb{E}\left[A_\psi^2 \mathbb{1}_{B_\varepsilon(b)}\right] \\
&\lesssim \frac{1}{c}(mb^{2\beta+2} \wedge 1)\Big(1 + \int_{-1/b}^{1/b}\Big((1+|u|)^{-1}\,\mathbb{E}[\Delta_m^2(u) + \Delta_m^4(u)] \\
&\quad + \mathbb{E}\left[(1+\Delta_m^4(u))(1+|u|)^{2\beta+1}|\varphi_{\varepsilon,m}'(u) - \varphi_\varepsilon'(u)|^2\right]\Big)\,\mathrm{d}u\Big) \\
&\lesssim \frac{mb^{2\beta+2} \wedge 1}{c}\Big(1 + \frac{1}{|\log b|^3}\int_{-1/b}^{1/b}\frac{\mathrm{d}u}{1+|u|} + \frac{1}{m}\int_{-1/b}^{1/b}(1+|u|)^{2\beta+1}\,\mathrm{d}u\Big) \lesssim \frac{1}{c}, \qquad (24)
\end{aligned}
$$

which completes the proof. $\qquad\square$

A.2. Proof of Proposition 2.2

The following lemma establishes a bound for the bias term of the estimator for the distribution function.

Lemma A.2. *Let Assumption 1 hold with $\ell = \langle\alpha\rangle + 1$, $\alpha > 0$ and $f(\bullet + q_\tau) \in C^\alpha([-\zeta, \zeta], R)$. Then we have*

$$\sup_{f(\bullet+q_\tau)\in C^\alpha([-\zeta,\zeta],R)}\Big|\int_{-\infty}^{q_\tau} K_b * f(x)\,\mathrm{d}x - \int_{-\infty}^{q_\tau} f(x)\,\mathrm{d}x\Big| \leqslant Db^{\alpha+1},$$

where $D = (R/(\langle\alpha\rangle + 1)! + 2\zeta^{-\alpha-1})\|K(x)x^{\alpha+1}\|_{L^1}$.

Proof. Let $F(x) := \int_{-\infty}^x f(y)\,\mathrm{d}y$. Twofold application of Fubini's theorem yields

$$
\begin{aligned}
\int_{-\infty}^{q_\tau} K_b * f(x)\,\mathrm{d}x &= \int_{-\infty}^{q_\tau}\int_{\mathbb{R}} \mathcal{F}^{-1}\Big[\frac{\varphi_K(bu)}{\varphi_\varepsilon(u)}e^{iuy}\Big](x)f_Y(y)\,\mathrm{d}y\,\mathrm{d}x \\
&= \int_{-\infty}^{\infty} K_b(x)F(q_\tau - x)\,\mathrm{d}x, \qquad (25)
\end{aligned}
$$

where $K_b(x) := b^{-1}K(x/b), x \in \mathbb{R}$. Therefore, the bias depends only locally on f. Note that $F(\bullet + q_\tau) \in C^{\alpha+1}([-\zeta, \zeta])$ by assumption. A Taylor expansion of F around q_τ yields for $|bz| < \zeta$

$$F(q_\tau - bz) - F(q_\tau) = -bzF'(q_\tau) + \cdots + (-bz)^{\langle\alpha\rangle+1}\frac{F^{(\langle\alpha\rangle+1)}(q_\tau - \kappa bz)}{(\langle\alpha\rangle + 1)!},$$

where $0 \leqslant \kappa \leqslant 1$. Using the fact that $\int x^k K(x)\,\mathrm{d}x = 0$ for $k = 1, \ldots, \langle\alpha\rangle + 1$ and the properties of the class, we obtain

$$
\begin{aligned}
\Big|\int_{-\infty}^{q_\tau}\big(K_b * f(x) - f(x)\big)\,\mathrm{d}x\Big| &= \Big|\int_{-\infty}^{\infty} K(z)\big(F(q_\tau - bz) - F(q_\tau)\big)\,\mathrm{d}z\Big| \\
&= \Big|\int_{|z|<\zeta/b} K(z)(-bz)^{\langle\alpha\rangle+1}\frac{F^{(\langle\alpha\rangle+1)}(q_\tau - \kappa bz) - F^{(\langle\alpha\rangle+1)}(q_\tau)}{(\langle\alpha\rangle + 1)!}\,\mathrm{d}z\Big|
\end{aligned}
$$

$$+ \int_{|z| \geqslant \zeta/b} |K(z)| |F(q_\tau - bz) - F(q_\tau)| \, \mathrm{d}z$$

$$\leqslant \frac{b^{\langle \alpha \rangle + 1} R}{(\langle \alpha \rangle + 1)!} \int_{-\infty}^{\infty} |K(z)| |z|^{\langle \alpha \rangle + 1} |\kappa bz|^{\alpha + 1 - (\langle \alpha \rangle + 1)} \, \mathrm{d}z + 2 \int_{|z| \geqslant \zeta/b} |K_b(z)| \, \mathrm{d}z$$

$$\leqslant \Big(\frac{b^{\alpha + 1} R}{(\langle \alpha \rangle + 1)!} + 2 \Big(\frac{b}{\zeta} \Big)^{\alpha + 1} \Big) \int_{-\infty}^{\infty} |K(z)| |z|^{\alpha + 1} \, \mathrm{d}z,$$

and the statement follows. $\qquad\square$

Proof of Proposition 2.2. We will show uniformly over $f \in \mathcal{C}^\alpha(R, r, \zeta)$ and $f_\varepsilon \in \mathcal{D}^\beta(R, \gamma)$ for any b such that $(n \wedge m) b^{2\beta + 1} \to \infty$

$$\Big| \int_{-\infty}^{q_\tau} (\widetilde{f}_b(x) - f(x)) \, \mathrm{d}x \Big| = \mathcal{O}_P \Big(b^{\alpha + 1} + \frac{1}{\sqrt{(n \wedge m)(b^{2\beta - 1} \wedge 1)}} + \frac{1}{\sqrt{(n \wedge m)(mb^{2\beta + 2} \wedge 1)}} \Big).$$

The third term on the right-hand side is of smaller or of the same order than the second one if and only if $(mb^{1 \wedge 2\beta + 2})^{-1} \lesssim 1$. Hence, when $\alpha \geqslant 1/2$ the asymptotically optimal choice $b = (n \wedge m)^{-1/(2\alpha + 2(\beta \vee 1/2) + 1)}$ yields

$$\Big| \int_{-\infty}^{q_\tau} (\widetilde{f}_b(x) - f(x)) \, \mathrm{d}x \Big| = \mathcal{O}_P \Big((n \wedge m)^{-(\alpha + 1)/(2\alpha + 2\beta + 1)} \vee (n \wedge m)^{-1/2} \Big).$$

Step 1: As usual we decompose the error into a deterministic error term and a stochastic error term, writting $\varphi_X = \mathcal{F} f$,

$$\Big| \int_{-\infty}^{q_\tau} \widetilde{f}_b(x) - f(x) \, \mathrm{d}x \Big| \leqslant \Big| \int_{-\infty}^{q_\tau} K_b * f(x) - f(x) \, \mathrm{d}x \Big|$$

$$+ \Big| \int_{-\infty}^{q_\tau} \mathcal{F}^{-1} \Big[\frac{\varphi_n(u) \varphi_K(bu)}{\varphi_{\varepsilon,m}(u)} - \varphi_K(bu) \varphi_X(u) \Big](x) \, \mathrm{d}x \Big|.$$

The bias is of order $\mathcal{O}(b^{\alpha + 1})$ by Lemma A.2. As discussed above, we decompose the stochastic error into a singular part and a continuous one using a smooth truncation function. Let $a_c \in C^\infty(\mathbb{R})$ satisfy $a_c(x) = 1$ for $x \leqslant -1$ and $a_c(x) = 0$ for $x \geqslant 0$ and define $a_s(x) := \mathbb{1}_{(-\infty, 0]}(x) - a_c(x)$. Then

$$\int_{-\infty}^{q_\tau} \mathcal{F}^{-1} \Big[\varphi_K(bu) \Big(\frac{\varphi_n(u)}{\varphi_{\varepsilon,m}(u)} - \varphi_X(u) \Big) \Big](x) \, \mathrm{d}x$$

$$= \int_{\mathbb{R}} a_s(x) \, \mathcal{F}^{-1} \Big[\varphi_K(bu) \Big(\frac{\varphi_n(u)}{\varphi_{\varepsilon,m}(u)} - \varphi_X(u) \Big) \Big](x + q_\tau) \, \mathrm{d}x$$

$$+ \int_{\mathbb{R}} a_c(x) \, \mathcal{F}^{-1} \Big[\varphi_K(bu) \Big(\frac{\varphi_n(u)}{\varphi_{\varepsilon,m}(u)} - \varphi_X(u) \Big) \Big](x + q_\tau) \, \mathrm{d}x =: T_s + T_c. \qquad (26)$$

The singular term T_s will be treated in the next step while we bound the continuous, but not integrable term T_c in Step 3.

Step 2: Recalling the definition (18) of the event $B_\varepsilon(b)$, let us denote its complement by $B_\varepsilon(b)^c$. Lemma A.1 shows that $B_\varepsilon(b)^c$ is asymptotically a null set. We obtain for any $c > 0$ with Markov's inequality

$$P \Big(|T_s| > \frac{c}{\sqrt{(n \wedge m)(b^{2\beta - 1} \vee 1)}} \Big) \leqslant P \Big(B_\varepsilon(b) \cap \Big\{ |T_s| > \frac{c}{\sqrt{(n \wedge m)(b^{2\beta - 1} \vee 1)}} \Big\} \Big) + P \Big(B_\varepsilon(b)^c \Big)$$

$$\leqslant \frac{1}{c} \sqrt{(n \wedge m)(b^{2\beta - 1} \vee 1)} \, \mathbb{E} \Big[|T_s| \mathbb{1}_{B_\varepsilon(b)} \Big] + o(1).$$

To bound $\mathbb{E}[|T_s| \mathbb{1}_{B_\varepsilon(b)}]$, we first note by Plancherel's identity

$$T_s = \frac{1}{2\pi} \int_{\mathbb{R}} \mathcal{F} a_s(u) e^{-iuq_\tau} \varphi_K(bu) \Big(\frac{\varphi_n(u)}{\varphi_{\varepsilon,m}(u)} - \varphi_X(u) \Big) \, \mathrm{d}u$$

$$= \frac{1}{2\pi} \int_{\mathbb{R}} \mathcal{F} a_s(u) e^{-iuq_\tau} \varphi_K(bu) \Big(\frac{\varphi_n(u)}{\varphi_\varepsilon(u)} - \varphi_X(u) \Big) \, du$$

$$+ \frac{1}{2\pi} \int_{\mathbb{R}} \mathcal{F} a_s(u) e^{-iuq_\tau} \frac{\varphi_K(bu)\varphi_n(u)}{\varphi_\varepsilon(u)} \Big(\frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)} - 1 \Big) \, du =: \frac{1}{2\pi} \big(T_{s,x} + T_{s,\varepsilon} \big). \tag{27}$$

The first term $T_{s,x}$ corresponds to the error due to the unknown density f while $T_{s,\varepsilon}$ is dominated by error of the estimator $\varphi_{\varepsilon,m}$. Since a_s is of bounded variation and has compact support, there is some constant $A_s \in (0, \infty)$ such that $|\mathcal{F} a_s(u)| \leqslant A_s(1 + |u|)^{-1}$. We then obtain with Plancherel's identity

$$\begin{aligned} \mathrm{Var}(T_{s,x}) = \mathbb{E}[|T_{s,x}|^2] &\leqslant \frac{1}{n} \mathbb{E} \Big[\Big| \int_{\mathbb{R}} \mathcal{F} a_s(u) e^{-iuq_\tau} \frac{\varphi_K(bu)}{\varphi_\varepsilon(u)} e^{iuY_1} \, du \Big|^2 \Big] \\ &\leqslant \frac{2\pi}{n} \|f_Y\|_\infty \Big\| \mathcal{F}^{-1} \Big[\mathcal{F} a_s(u) \frac{\varphi_K(bu)}{\varphi_\varepsilon(u)} \Big] \Big\|_{L^2}^2 \\ &\leqslant \frac{1}{n} \|K\|_{L^1}^2 \|f_Y\|_\infty \int_{-1/b}^{1/b} \frac{|\mathcal{F} a_s(u)|^2}{|\varphi_\varepsilon(u)|^2} \, du \\ &\leqslant \frac{1}{n} \|K\|_{L^1}^2 A_s^2 \|f_Y\|_\infty \int_{-1/b}^{1/b} \frac{1}{(1 + |u|)^2 |\varphi_\varepsilon(u)|^2} \, du. \end{aligned}$$

Using the assumption $\|f\|_\infty < R$ and $f_\varepsilon \in \mathcal{D}^\beta(R, \gamma)$, we get

$$\mathbb{E}[|T_{s,x}|^2] \lesssim \frac{1}{n} \int_{-1/b}^{1/b} (1 + |u|)^{2\beta-2} \, du \lesssim \frac{1}{nb^{2\beta-1}} \vee \frac{1}{n}. \tag{28}$$

To bound $T_{s,\varepsilon}$, we will use the following version of a lemma by Neumann (1997): By the definition (18) of $B_\varepsilon(b)$ and applying (20) it holds

$$\begin{aligned} \mathbb{E} \Big[\Big| \frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)} - 1 \Big|^2 \mathbb{1}_{B_\varepsilon(b)} \Big] &\leqslant 2 \mathbb{E} \Big[\frac{|\varphi_{\varepsilon,m}(u) - \varphi_\varepsilon(u)|^2}{|\varphi_\varepsilon(u)|^2} \Big] + 2 \mathbb{E} \Big[\frac{|\varphi_{\varepsilon,m}(u) - \varphi_\varepsilon(u)|^4}{|\varphi_\varepsilon(u)\varphi_{\varepsilon,m}(u)|^2} \mathbb{1}_{B_\varepsilon(b)} \Big] \\ &\leqslant \frac{2\, \mathbb{E}[|\varphi_{\varepsilon,m}(u) - \varphi_\varepsilon(u)|^2]}{|\varphi_\varepsilon(u)|^2} + \frac{2m\, \mathbb{E}[|\varphi_{\varepsilon,m}(u) - \varphi_\varepsilon(u)|^4]}{|\varphi_\varepsilon(u)|^2} \\ &\leqslant \frac{18}{m|\varphi_\varepsilon(u)|^2}. \end{aligned} \tag{29}$$

Now, we estimate with the Cauchy–Schwarz inequality

$$\begin{aligned} T_{s,\varepsilon}^2 &\leqslant \|K\|_{L^1}^2 \int_{-1/b}^{1/b} \Big| \frac{\varphi_n(u)}{\varphi_\varepsilon(u)} \Big|^2 \, du \int_{-1/b}^{1/b} |\mathcal{F} a_s(u)|^2 \Big| \frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)} - 1 \Big|^2 \, du \\ &\leqslant 2\|K\|_{L^1}^2 \Big(\|\varphi_X\|_{L^2}^2 + \int_{-1/b}^{1/b} \frac{|\varphi_n(u) - \varphi_Y(u)|^2}{|\varphi_\varepsilon(u)|^2} \, du \Big) \int_{-1/b}^{1/b} |\mathcal{F} a_s(u)|^2 \Big| \frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)} - 1 \Big|^2 \, du. \end{aligned}$$

Applying again the Cauchy–Schwarz inequality, Fubini's theorem, the decay of $\mathcal{F} a_s$ and (29), we obtain

$$\begin{aligned} \mathbb{E}[|T_{s,\varepsilon}| \mathbb{1}_{B_\varepsilon(b)}] &\leqslant \sqrt{2} \|K\|_{L^1} \Big(\|\varphi_X\|_{L^2}^2 + \int_{-1/b}^{1/b} \frac{\mathbb{E}[|\varphi_n(u) - \varphi_Y(u)|^2]}{|\varphi_\varepsilon(u)|^2} \, du \Big)^{1/2} \\ &\quad \times \Big(\int_{-1/b}^{1/b} \frac{A_s^2}{(1 + |u|)^2} \mathbb{E} \Big[\Big| \frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)} - 1 \Big|^2 \mathbb{1}_{B_\varepsilon(b)} \Big] \, du \Big)^{1/2} \\ &\leqslant \frac{\sqrt{36} \|K\|_{L^1} A_s}{\sqrt{m}} \Big(\|\varphi_X\|_{L^2}^2 + \int_{-1/b}^{1/b} \frac{1}{n|\varphi_\varepsilon(u)|^2} \, du \Big)^{1/2} \\ &\quad \times \Big(\int_{-1/b}^{1/b} \frac{1}{(1 + |u|)^2 |\varphi_\varepsilon(u)|^2} \, du \Big)^{1/2}. \end{aligned} \tag{30}$$

The assumptions $\|f\|_\infty \lesssim 1$ and $|\varphi_\varepsilon(u)| \lesssim (1 + |u|)^{-\beta}$ yield

$$\mathbb{E}[|T_{s,\varepsilon}|\mathbb{1}_{B_\varepsilon(b)}] \lesssim \Big(1 + \frac{1}{nb^{2\beta+1}}\Big)^{1/2}\Big(\frac{1}{\sqrt{m}b^{\beta-1/2}} \vee \frac{1}{\sqrt{m}}\Big) \lesssim \frac{1}{\sqrt{m}b^{\beta-1/2}} \vee \frac{1}{\sqrt{m}}.$$

The last estimate follows from the fact that the choice of b as stated in the theorem guarantees that $n^{-1}b^{-2\beta-1} \to 0$ as $n \to \infty$. The last inequality together with (28) and (27) imply the optimal order

$$\mathbb{E}[|T_s|\mathbb{1}_{B_\varepsilon(b)}] \lesssim \Big((n \wedge m)(b^{2\beta-1} \wedge 1)\Big)^{-1/2}.$$

Step 3: Let us define the empirical measures of (Y_j) and (ε_k) as $\mu_{Y,n} := \frac{1}{n}\sum_{j=1}^n \delta_{Y_j}$ and $\mu_{\varepsilon,m} := \frac{1}{m}\sum_{k=1}^m \delta_{\varepsilon_k}$, respectively, where δ_x is the Dirac measure in $x \in \mathbb{R}$. We can write

$$\begin{aligned}
T_c &= \int_\mathbb{R} a_c(x)\,\mathcal{F}^{-1}\Big[\frac{\varphi_K(bu)}{\varphi_{\varepsilon,m}(u)}\Big(\varphi_n(u) - \varphi_{\varepsilon,m}(u)\varphi_X(u)\Big)\Big](x + q_\tau)\,\mathrm{d}x \\
&= \mathcal{F}^{-1}\Big[\frac{\varphi_K(-bu)}{\varphi_{\varepsilon,m}(-u)}\Big(\varphi_n(-u) - \varphi_{\varepsilon,m}(-u)\varphi_X(-u)\Big)\Big] * a_c(-q_\tau) \\
&= \mathcal{F}^{-1}\Big[\frac{\varphi_K(bu)}{\varphi_{\varepsilon,m}(u)}\Big] * \Big(\mu_{Y,n} * a_c(-\bullet) - \mu_{\varepsilon,m} * f * a_c(-\bullet)\Big)(q_\tau).
\end{aligned}$$

Applying Lemma 2.1, we obtain on $B_\varepsilon(b)$ for any integer $s > \beta$

$$\begin{aligned}
T_c &\leq \Big\|\mathcal{F}^{-1}\Big[\frac{\varphi_K(bu)}{\varphi_{\varepsilon,m}(u)}\Big] * \Big(\mu_{Y,n} * a_c(-\bullet) - \mu_{\varepsilon,m} * f * a_c(-\bullet)\Big)\Big\|_\infty \\
&\leq \mathcal{E}_b\Big\|\mu_{Y,n} * a_c(-\bullet) - \mu_{\varepsilon,m} * f * a_c(-\bullet)\Big\|_{C^s} \\
&= \mathcal{E}_b\sum_{l=0}^s\big\|\mu_{Y,n} * a_c^{(l)}(-\bullet) - \mu_{\varepsilon,m} * f * a_c^{(l)}(-\bullet)\big\|_\infty.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&P\Big(B_\varepsilon(b) \cap \Big\{|T_c| > \frac{c}{\sqrt{(n \wedge m)}(\sqrt{m}b^{\beta+1} \wedge 1)}\Big\}\Big) \\
&\leq P\Big(B_\varepsilon(b) \cap \Big\{\mathcal{E}_b > \Big(\frac{c}{mb^{2\beta+2} \wedge 1}\Big)^{1/2}\Big\}\Big) \\
&\quad + P\Big(\sum_{l=0}^s\big\|\mu_{Y,n} * a_c^{(l)} - \mu_{\varepsilon,m} * f * a_c^{(l)}\big\|_\infty > \Big(\frac{c}{n \wedge m}\Big)^{1/2}\Big) =: P_1 + P_2.
\end{aligned}$$

By Lemma 2.1, more precisely estimate (24), the first probability is of the order $1/c$. To bound P_2, is suffices to show $\big\|\mu_{Y,n} * a_c^{(l)} - \mu_{\varepsilon,m} * f * a_c^{(l)}\big\|_\infty = \mathcal{O}_P((n \wedge m)^{-1/2})$ for all $l = 0, \ldots, s$. Denoting the density of Y_j as $f_Y = f * f_\varepsilon$, we decompose

$$\begin{aligned}
&\big\|\mu_{Y,n} * (a_c^{(l)}(-\bullet)) - \mu_{\varepsilon,m} * f * (a_c^{(l)}(-\bullet))\big\|_\infty \\
&\leq \big\|\mu_{Y,n} * (a_c^{(l)}(-\bullet)) - f_Y * (a_c^{(l)}(-\bullet))\big\|_\infty + \big\|f_\varepsilon * (f * (a_c^{(l)}(-\bullet))) - \mu_{\varepsilon,m} * (f * (a_c^{(l)}(-\bullet)))\big\|_\infty \\
&\leq \Big\|\int a_c^{(l)}(y - \bullet)\mu_{Y,n}(\mathrm{d}y) - \mathbb{E}[a_c^{(l)}(Y_1 - \bullet)]\Big\|_\infty \\
&\quad + \Big\|\mathbb{E}[(f * a_c^{(l)})(\varepsilon_1 - \bullet)] - \int (f * a_c^{(l)})(z - \bullet)\mu_{\varepsilon,m}(\mathrm{d}z)\Big\|_\infty
\end{aligned}$$

By construction all $a_c^{(l)}, l \geq 1$, have compact support and are bounded. Therefore, $\|a_c^{(l)}\|_{L^1} < \infty$ and $\|(a_c * f)^{(l)}\|_{L^1} \leq \|a_c^{(l)}\|_{L^1}\|f\|_{L^1} < \infty$ and thus the functions $a_c(\bullet - t)$ and $a_c * f(\bullet - t)$ for all $t \in \mathbb{R}$ are of bounded variation. Since the set of functions with bounded variation is a Donsker class

(cf. Theorem 2.1 by Dudley (1992)), the two terms in the previous display converge in probability to a tight limit with $\sqrt{n}$-rate and $\sqrt{m}$-rate, respectively. Consequently,

$$\sqrt{n \wedge m} \big\| \mu_{Y,n} * (a_c^{(l)}(-\bullet)) - \mu_{\varepsilon,m} * f * (a_c^{(l)}(-\bullet)) \big\|_\infty = \mathcal{O}_P(1)$$

for all $\ell = 0, \ldots, s$ and P_2 is arbitrary small for c large. $\qquad \square$

For the adaptive estimator we will later need the following uniform version of Proposition 2.2.

Corollary A.3. *Suppose Assumption 1 holds with $l = \langle \alpha \rangle + 1$ and let $\mathcal{B}$ be a finite set of bandwidths with $b_1 = \min \mathcal{B}$ such that $m b_1^{2\beta_1 + 1} \to \infty$. For a sequence of critical values $(\delta_b)_{b \in \mathcal{B}}$ satisfying $\delta_b > 3 D b^{\alpha+1}$ and for any sequence $(x_n)_n$ with $x_n \to \infty$ arbitrary slowly we obtain uniformly in $\mathcal{C}^\alpha(R, r, \zeta)$ and $\mathcal{D}^\beta(R, \gamma)$*

$$P\Big(\exists b \in \mathcal{B} : \Big| \int_{-\infty}^{q_\tau} \big(\widetilde{f}_b(x) - f(x) \big) \, \mathrm{d}x \Big| > \delta_b \Big)$$
$$= \mathcal{O}\Big(\sum_{b \in \mathcal{B}} \Big(\frac{1}{\delta_b} \big((n \wedge m)(b^{2\beta-1} \wedge 1) \big)^{-1/2} + \frac{1}{\delta_b^2} \frac{x_n}{n(m b^{2\beta+2} \wedge 1)} \Big) \Big) + o(1).$$

In particular, if $|\mathcal{B}| \lesssim \log n$, $\max_{b \in \mathcal{B}} b \to 0$ and $\min_{b \in \mathcal{B}}(n \wedge m) b^{2\beta+1} \to \infty$

$$P\Big(\sup_{b \in \mathcal{B}} \Big| \int_{-\infty}^{q_\tau} \widetilde{f}_b(x) - f(x) \, \mathrm{d}x \Big| > \delta \Big) \to 0 \quad \text{for all } \delta > 0.$$

Proof. With the notation of the proof of Proposition 2.2 and applying Lemma A.2, we obtain

$$\Big| \int_{-\infty}^{q_\tau} \big(\widetilde{f}_b(x) - f(x) \big) \, \mathrm{d}x \Big| \leqslant \Big| \int_{-\infty}^{q_\tau} \big(K_b * f(x) - f(x) \big) \, \mathrm{d}x \Big| + T_s + T_c \leqslant D b^{\alpha+1} + T_s + T_c,$$

where T_s and T_c are the stochastic errors of the singular part and of the continuous part, respectively, as defined in (26). Since both terms depend on b let us write $T_s(b)$ and $T_c(b)$. By definition $b_1 \leqslant b$ implies $B_\varepsilon(b_1) \subseteq B_\varepsilon(b)$. Then, Step 2 in the previous proof shows

$$\begin{aligned}
P\big(\exists b \in \mathcal{B} : T_s > \delta_b/3 \big) &\leqslant P\big(\exists b \in \mathcal{B} : T_s(b) > \delta_b/3 \} \cap B_\varepsilon(b_1) \big) + P\big(B_\varepsilon(b_1)^c \big) \\
&\leqslant \Big(\sum_{b \in \mathcal{B}} P\big(\{ T_s(b) > \delta_b/3 \} \cap B_\varepsilon(b_1) \big) \Big) + o(1) \\
&\leqslant \Big(\sum_{b \in \mathcal{B}} \delta_b^{-1} \, \mathbb{E}[|T_s(b)| \mathbb{1}_{B_\varepsilon(b_1)}] \Big) + o(1) \\
&\lesssim \Big(\sum_{b \in \mathcal{B}} \delta_b^{-1} \big((n \wedge m)(b^{2\beta-1} \wedge 1) \big)^{-1/2} \Big) + o(1).
\end{aligned}$$

Following Step 3 in the previous proof, we obtain with the random operator norm $\mathcal{E}_b$, for some integer $s > \beta$ and for a diverging sequence (x_n)

$$\begin{aligned}
P\big(\exists b \in \mathcal{B} : T_c > \delta_b/3 \big) &\leqslant P\big(\{ \exists b \in \mathcal{B} : \mathcal{E}_b > \delta_b n^{1/2}/(3(x_n)^{1/2}) \} \cap B_\varepsilon(b_1) \big) + P\big(B_\varepsilon(b_1)^c \big) \\
&\quad + P\Big(\Big\{ \sum_{l=0}^s \big\| \mu_{Y,n} * a_c^{(l)} - \mu_{\varepsilon,m} * f * a_c^{(l)} \big\|_\infty > \Big(\frac{x_n}{n} \Big)^{1/2} \Big\} \Big) \\
&\leqslant \Big(\sum_{b \in \mathcal{B}} P\big(\{ \mathcal{E}_b > \delta_b n^{1/2}/(3(x_n)^{1/2}) \} \cap B_\varepsilon(b_1) \big) \Big) + o(1) \\
&\lesssim \Big(\sum_{b \in \mathcal{B}} \frac{x_n}{\delta_b^2 n(m b^{2\beta+2} \wedge 1)} \Big) + o(1),
\end{aligned}$$

where we have used (24) in the last estimate. $\qquad \square$

A.3. Proof of Proposition 2.3

Without loss of generality we set $q_\tau = 0$. Recall definition (11) of the pseudo-estimator $\widehat{f}_b$ which knows the error distribution. We estimate

$$\sup_{x \in (-\zeta, \zeta)} |\widetilde{f}_b(x) - f(x)| \leqslant \sup_{x \in (-\zeta, \zeta)} |\widehat{f}_b(x) - f(x)| + \|\widetilde{f}_b - \widehat{f}_b\|_\infty$$

$$\leqslant \sup_{x \in (-\zeta, \zeta)} |\widehat{f}_b(x) - f(x)| + \left\| \frac{\varphi_K(bu)\varphi_n(u)}{\varphi_\varepsilon(u)} \left(\frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)} - 1 \right) \right\|_{L^1}.$$

The analysis of the first term is very classical. However, the authors are not aware of any reference in the given setup. Both terms will be treated separately in the following two steps. All estimates will be uniform in $f \in \mathcal{C}^\alpha(R, r, \zeta)$ and $f_\varepsilon \in \mathcal{D}^\beta(R, \gamma)$.

Step 1: Let $b \in (0, 1)$. We will show that there are constants $d, D > 0$ such that for any $t > d(b^\alpha + (nb^{2\beta+1})^{-2})$

$$P\left(\sup_{x \in (-\zeta, \zeta)} |\widehat{f}_b(x) - f(x)| > t \right) \leqslant 2 \exp\left(2 \log n - D n b^{(2\beta+1)} (t \wedge t^2) \right). \tag{31}$$

Then the result follows by choosing $t \sim b^\alpha + \left(\frac{\log n}{nb^{2\beta+1}} \right)^{1/2}$. Let us define $x_k := -\zeta + kn^{-2}$ for $k = 1, \ldots, \lfloor 2\zeta n^2 \rfloor =: M$ as well as

$$\chi_j(x) := \mathcal{F}^{-1}\left[\frac{\varphi_K(bu)}{\varphi_\varepsilon(u)} e^{iuY_j} \right](x) - \mathbb{E}\left[\mathcal{F}^{-1}\left[\frac{\varphi_K(bu)}{\varphi_\varepsilon(u)} e^{iuY_j} \right](x) \right]$$

$$= K_b * \mathcal{F}^{-1}\left[\mathbb{1}_{[-b^{-1}, b^{-1}]}(u) \frac{e^{iuY_j}}{\varphi_\varepsilon(u)} \right](x) - K_b * f(x), \quad x \in \mathbb{R}.$$

Therefore, $\widehat{f}_b(x) - \mathbb{E}[\widehat{f}_b(x)] = \frac{1}{n} \sum_{j=1}^n \chi_j(x)$ and thus

$$\sup_{|x| < \zeta} |\widehat{f}_b(x) - f(x)| \leqslant \sup_{|x| < \zeta} |\mathbb{E}[\widehat{f}_b(x)] - f(x)| + \sup_{|x| < \zeta} |\widehat{f}_b(x) - \mathbb{E}[\widehat{f}_b(x)]|$$

$$\leqslant \sup_{|x| < \zeta} |\mathbb{E}[\widehat{f}_b(x)] - f(x)| + \sup_{|x| < \zeta} \min_{k=1,\ldots,M} \left| \frac{1}{n} \sum_{j=1}^n (\chi_j(x) - \chi_j(x_k)) \right|$$

$$+ \max_{k=1,\ldots,M} \left| \frac{1}{n} \sum_{j=1}^n \chi_j(x_k) \right|$$

$$=: B + V_1 + V_2.$$

The bias term B can be bounded as in the classical density estimation setup (cf. also Fan, 1991, Thm. 1 and 2), noting that the constant does not depend on $x \in (-\zeta, \zeta)$. Hence, $|B| \lesssim b^\alpha$. Using a continuity argument and the properties of $f_\varepsilon \in \mathcal{D}^\beta(R, \gamma)$, the term V_1 can be bounded by

$$|V_1| \leqslant \frac{1}{n^2} \left\| \frac{1}{n} \sum_{j=1}^n \chi_j' \right\|_\infty = \frac{1}{n^3} \left\| \sum_{j=1}^n (K_b') * \left(\mathcal{F}^{-1}\left[\mathbb{1}_{[-b^{-1}, b^{-1}]}(u) \frac{e^{iuY_j}}{\varphi_\varepsilon(u)} \right] - f \right) \right\|_\infty$$

$$\leqslant \frac{1}{n^2 b} \|K'\|_{L^1} \left(\|\mathbb{1}_{[-b^{-1}, b^{-1}]} \varphi_\varepsilon^{-1}\|_{L^1} + \|f\|_\infty \right) \lesssim n^{-2} b^{-(\beta+2)} \lesssim (nb^{2\beta+1})^{-2}.$$

Therefore, $|B + V_1| \leqslant D_1(b^\alpha + (nb^{2\beta+1})^{-2})$ for some constant $D_1 > 0$. We obtain for all $t > d(b^\alpha + (nb^{2\beta+1})^{-2})$ with $d := 2D_1$

$$P\left(\sup_{|x| < \zeta} |\widehat{f}_b(x) - f(x)| > t \right) \leqslant P\left(\max_{k=1,\ldots,M} \left| \frac{1}{n} \sum_{j=1}^n \chi_j(x_k) \right| > \frac{t}{2} \right) \leqslant \sum_{k=1}^M P\left(\left| \frac{1}{n} \sum_{j=1}^n \chi_j(x_k) \right| > \frac{t}{2} \right).$$

Finally, we will apply Bernstein's inequality. To this end we estimate

$$\max_{j,k} |\chi_j(x_k)| \leqslant 2\|K_b\|_{L^1} \|\mathbb{1}_{[-b^{-1},b^{-1}]}\varphi_\varepsilon^{-1}\|_{L^1} \leqslant D_2 b^{-(\beta+1)},$$

with some constant $D_2 > 0$. Using Plancherel's identity, the variance can be estimated by

$$\mathrm{Var}(\chi_j(x_k)) = \mathbb{E}\left[\mathcal{F}^{-1}\left[\frac{\varphi_K(bu)}{\varphi_\varepsilon(u)}e^{iuY_j}\right]^2(x_k)\right] - (K_b * f)^2(x_k)$$
$$\leqslant \frac{1}{2\pi}\|f\|_\infty \left\|\frac{\varphi_K(-bu)}{\varphi_\varepsilon(-u)}\right\|_{L^2}^2 \lesssim D_3 b^{-(2\beta+1)},$$

for some $D_3 > 0$. Then Bernstein's inequality yields

$$P\left(\sup_{x\in(-\zeta,\zeta)}|\widehat{f}_b(x) - f(x)| > t\right) \leqslant \sum_{k=1}^{M} P\left(\left|\sum_{j=1}^{n}\chi_j(x_k)\right| > nt/2\right)$$
$$\leqslant 2M\exp\left(-\frac{nt^2}{8D_3 b^{-(2\beta+1)} + \frac{8}{3}D_2 b^{-(\beta+1)}t}\right)$$
$$\leqslant 2\exp\left(\log M - \frac{nb^{(2\beta+1)}t^2}{8(D_3 + D_2 t/3)}\right)$$
$$\leqslant 2\exp\left(2\log n - Dnb^{(2\beta+1)}(t\wedge t^2)\right),$$

with some constant $D > 0$.

Step 2: By the Cauchy–Schwarz inequality we have

$$\mathbb{E}\left[\left\|\frac{\varphi_K(bu)\varphi_n(u)}{\varphi_\varepsilon(u)}\left(\frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)} - 1\right)\right\|_{L^1}\mathbb{1}_{B_\varepsilon(b)}\right]$$
$$\lesssim \left(\mathbb{E}\left[\left\|\frac{\varphi_n(u)}{\varphi_\varepsilon(u)}\mathbb{1}_{[-1/b,1/b]}(u)\right\|_{L^2}^2\right]\mathbb{E}\left[\left\|\left(\frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)} - 1\right)\mathbb{1}_{[-1/b,1/b]}(u)\right\|_{L^2}^2\mathbb{1}_{B_\varepsilon(b)}\right]\right)^{1/2}$$
$$\leqslant \left(\|\varphi_X\|_{L^2} + \left(\int_{-1/b}^{1/b}\frac{\mathbb{E}[|\varphi_n(u) - \varphi_Y(u)|^2]}{|\varphi_\varepsilon(u)|^2}\,du\right)^{1/2}\right)\left(\int_{-1/b}^{1/b}\mathbb{E}\left[\left|\frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)} - 1\right|^2\mathbb{1}_{B_\varepsilon(b)}\right]\,du\right)^{1/2}$$
$$\lesssim \left(\|\varphi_X\|_{L^2} + \left(\frac{1}{nb^{2\beta+1}}\right)^{1/2}\right)\left(\frac{1}{mb^{2\beta+1}}\right)^{1/2},$$

where we have used (29) for the last step. Therefore, the additional error due to the unknown error distribution satisfies for any $\delta > 0$ by Markov's inequality and by Lemma A.1

$$P\left(\left\|\frac{\varphi_K(bu)\varphi_n(u)}{\varphi_\varepsilon(u)}\left(\frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)} - 1\right)\right\|_{L^1} > \delta\right) \leqslant \frac{1}{\delta}\mathbb{E}\left[\left\|\frac{\varphi_K(bu)\varphi_n(u)}{\varphi_\varepsilon(u)}\left(\frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)} - 1\right)\right\|_{L^1}\mathbb{1}_{B_\varepsilon(b)}\right]$$
$$+ P\left(\inf_{|u|\leqslant 1/b}|\varphi_{\varepsilon,m}(u)| < m^{-1/2}\right)$$
$$\lesssim \frac{1}{\delta}\left(\frac{1}{nb^{2\beta+1}}\right)^{1/2} + o(1). \tag{32}$$

and thus $\|\widetilde{f}_b - \widehat{f}_b\|_\infty = \mathcal{O}_P((mb^{2\beta+1})^{-1/2})$. Note that the second term does not depend on δ and thus $o(1)$ is sufficient. □

A.4. Proof of Theorem 2.4

We start with a lemma that establishes consistency of the quantile estimator and then prove the theorem. To apply this lemma also for the adaptive result, we prove convergence uniformly over a set of bandwidths.

Lemma A.4. *Grant Assumption 1 with $\ell = 1$. Let $\mathcal{B}$ be a set of bandwidths satisfying $|\mathcal{B}| \lesssim \log n$, $\max \mathcal{B} \to 0$ and $\min_{b \in \mathcal{B}} (\log n)^2/((n \wedge m) b^{2\beta+1}) \to 0$. Then*

$$\sup_{f \in \mathcal{C}^\alpha(R, r, \zeta)} \sup_{f_\varepsilon \in \mathcal{D}^\beta(R, \gamma)} P\big(\sup_{b \in \mathcal{B}} |\widetilde{q}_{\tau, b} - q_\tau| > \delta \big) \to 0 \quad \text{for all } \delta > 0.$$

Proof. We follow the general strategy of the proof of Theorem 5.7 by van der Vaart (1998) in the classical M-estimation setting. First, we show that f satisfies the uniqueness condition $\inf_{\eta: |\eta - q_\tau| \geqslant \delta} |M'(\eta)| > 0$ for any $\delta > 0$ and with $M'(\eta) = \int_{-\infty}^\eta f(x) \, dx - \tau$, which is the deterministic counterpart of $\widetilde{M}'_b$ defined in (7). By the Hölder regularity $M''(\eta) = f(\eta) \geqslant f(q_\tau) - |f(q_\tau) - f(\eta)| \geqslant r - R|q_\tau - \eta|^{1 \wedge \alpha} \geqslant r/2$ for $|q_\tau - \eta| \leqslant (\frac{r}{2R})^{1 \vee \alpha^{-1}}$. Without loss of generality we can assume $\delta \leqslant (\frac{r}{2R})^{1 \vee \alpha^{-1}}$, otherwise consider the minimum $\delta \wedge (\frac{r}{2R})^{1 \vee \alpha^{-1}}$. Recall that the true τ-quantile q_τ is given by the root of M' and that M' is monotone increasing. Hence, we obtain

$$\inf_{\eta: |\eta - q_\tau| \geqslant \delta} |M'(\eta)| = \inf_{\eta \in \{-\delta, \delta\}} \big| M'(q_\tau - \eta) - M'(q_\tau) \big| \geqslant \delta \inf_{\eta: |\eta - q_\tau| \geqslant \delta} M''(\eta) \geqslant \frac{\delta r}{2}.$$

Recall that $\widetilde{q}_{\tau, b}$ is given as the root of the estimating equation (7) on the interval $[-U_n, U_n]$ for $U_n \lesssim \log n$. Therefore,

$$
\begin{aligned}
P\big(\sup_{b \in \mathcal{B}} |\widetilde{q}_{\tau, b} - q_\tau| > \delta \big) &\leqslant P\big(\sup_{b \in \mathcal{B}} |M'(\widetilde{q}_{\tau, b})| \geqslant \delta r/2 \big) \\
&= P\big(\sup_{b \in \mathcal{B}} |M'(\widetilde{q}_{\tau, b}) - \widetilde{M}'_b(\widetilde{q}_{\tau, b})| \geqslant \delta r/2 \big) \\
&\leqslant P\big(\sup_{b \in \mathcal{B}} \sup_{\eta \in [-U_n, U_n]} |M'(\eta) - \widetilde{M}'_b(\eta)| \geqslant \delta r/2 \big) \\
&= P\Big(\sup_{b \in \mathcal{B}} \sup_{\eta \in [-U_n, U_n]} \Big| \int_{-\infty}^\eta (\widetilde{f}_b(x) - f(x)) \, dx \Big| \geqslant \delta r/2 \Big), \quad (33)
\end{aligned}
$$

where we have used $\widetilde{M}'_b(\widetilde{q}_{\tau, b}) = 0$. Hence, it remains to show uniform consistency of $\int_{-\infty}^\eta \widetilde{f}_b(x) \, dx$. Write,

$$
\begin{aligned}
\Big| \int_{-\infty}^\eta (\widetilde{f}_b(x) - f(x)) \, dx \Big| &\leqslant \Big| \int_{-\infty}^\eta (K_b * f(x) - f(x)) \, dx \Big| + \Big| \int_{-\infty}^\eta (\widetilde{f}_b(x) - K_b * f(x)) \, dx \Big| \\
&= |K_b * F(\eta) - F(\eta)| + \Big| \int_{-\infty}^\eta (\widetilde{f}_b(x) - K_b * f(x)) \, dx \Big|.
\end{aligned}
$$

We have $|K_b * F(\eta) - F(\eta)| = |\int K_b(z)(F(\eta - z) - F(\eta)) \, dz| \leqslant b \|f\|_\infty \|z K(z)\|_{L^1}$ by the boundedness of f. Further note for $\eta \in [-U_n, U_n]$

$$
\begin{aligned}
&\Big| \int_{-\infty}^\eta (\widetilde{f}_b(x) - K_b * f(x)) \, dx \Big| \\
&\leqslant \Big| \int_{-\infty}^{q_\tau} (\widetilde{f}_b(x) - K_b * f(x)) \, dx \Big| + \Big| \int_{q_\tau \wedge \eta}^{q_\tau \vee \eta} (\widetilde{f}_b(x) - K_b * f(x)) \, dx \Big| \\
&\leqslant \Big| \int_{-\infty}^{q_\tau} (\widetilde{f}_b(x) - K_b * f(x)) \, dx \Big| + \sqrt{2U_n} \Big(\int_{-\infty}^\infty (\widetilde{f}_b(x) - K_b * f(x))^2 \, dx \Big)^{1/2},
\end{aligned}
$$

where we have used the Cauchy-Schwarz inequality for the last step. Hence, together with (33) we obtain for all $\delta > 6 \|f\|_\infty \|z K(z)\|_{L^1}/r \sup_{b \in \mathcal{B}} b$

$$
\begin{aligned}
P\big(\sup_{b \in \mathcal{B}} |\widetilde{q}_{\tau, b} - q_\tau| > \delta \big) &\leqslant P\Big(\sup_{b \in \mathcal{B}} \sup_{\eta \in [-U_n, U_n]} \Big| \int_{-\infty}^\eta (\widetilde{f}_b(x) - f(x)) \, dx \Big| \geqslant \delta r/2 \Big) \\
&\leqslant P\Big(\sup_{b \in \mathcal{B}} \Big| \int_{-\infty}^{q_\tau} (\widetilde{f}_b(x) - K_b * f(x)) \, dx \Big| \geqslant \delta r/6 \Big)
\end{aligned}
$$

$$+ P\Big(\sup_{b \in \mathcal{B}} \int_{\mathbb{R}} (\widetilde{f}_b(x) - K_b * f(x))^2 \, \mathrm{d}x \geqslant \delta^2 r^2/(72 U_n) \Big).$$

Corollary A.3 shows under the conditions on $\mathcal{B}$ that

$$P\Big(\sup_{b \in \mathcal{B}} \Big| \int_{-\infty}^{q_\tau} \widetilde{f}_b(x) - K_b * f(x) \, \mathrm{d}x \Big| > \delta r/3 \Big) \to 0$$

Hence, it remains to show

$$P\Big(\sup_{b \in \mathcal{B}} \int_{\mathbb{R}} (\widetilde{f}_b(x) - K_b * f(x))^2 \, \mathrm{d}x > \delta^2 r^2/(72 U_n) \Big) \to 0. \tag{34}$$

On the event $B_\varepsilon(b)$, (34) follows basically from the work of Neumann (1997). More precisely, Plancherel's equality, (29) and the Cauchy–Schwarz inequality yield for any $b \in \mathcal{B}$

$$\mathbb{E} \Big[\int_{\mathbb{R}} (\widetilde{f}_b(x) - K_b * f(x))^2 \, \mathrm{d}x \mathbb{1}_{B_\varepsilon(b)} \Big]$$

$$= \frac{1}{2\pi} \int_{\mathbb{R}} |\varphi_K(bu)|^2 \, \mathbb{E} \Big[\Big| \frac{\varphi_n(u)}{\varphi_{\varepsilon,m}(u)} - \frac{\varphi_Y(u)}{\varphi_\varepsilon(u)} \Big|^2 \mathbb{1}_{B_\varepsilon(b)} \Big] \, \mathrm{d}u$$

$$\lesssim \int_{-1/b}^{1/b} \Big(\mathbb{E} \Big[\frac{|\varphi_n(u) - \varphi_Y(u)|^2}{|\varphi_{\varepsilon,m}(u)|^2} \mathbb{1}_{B_\varepsilon(b)} \Big] + |\varphi_Y(u)|^2 \, \mathbb{E} \Big[\Big| \frac{1}{\varphi_{\varepsilon,m}(u)} - \frac{1}{\varphi_\varepsilon(u)} \Big|^2 \mathbb{1}_{B_\varepsilon(b)} \Big] \Big) \, \mathrm{d}u$$

$$\lesssim \int_{-1/b}^{1/b} \Big(\mathbb{E} \Big[\frac{|\varphi_n(u) - \varphi_Y(u)|^2}{|\varphi_\varepsilon(u)|^2} \big(1 + m|\varphi_{\varepsilon,m}(u) - \varphi_\varepsilon(u)|^2 \big) \Big] + \frac{|\varphi_Y(u)|^2}{m|\varphi_\varepsilon(u)|^4} \Big) \, \mathrm{d}u$$

$$\leqslant \int_{-1/b}^{1/b} \frac{1}{|\varphi_\varepsilon(u)|^2} \Big(\Big(\mathbb{E} \big[|\varphi_n(u) - \varphi_Y(u)|^4 \big] \, \mathbb{E} \big[2 + 2m^2 |\varphi_{\varepsilon,m}(u) - \varphi_\varepsilon(u)|^4 \big] \Big)^{1/2} + \frac{|\varphi_X(u)|^2}{m} \Big) \, \mathrm{d}u$$

$$\lesssim \int_{-1/b}^{1/b} |\varphi_\varepsilon(u)|^{-2} (n^{-1} + m^{-1}) \, \mathrm{d}u \lesssim \frac{1}{(n \wedge m) b^{2\beta+1}}.$$

Noting that $B_\varepsilon(\min \mathcal{B}) \subseteq B_\varepsilon(b)$ and applying Lemma A.1, (34) follows then from Markov's inequality

$$P\Big(\sup_{b \in \mathcal{B}} \int_{\mathbb{R}} (\widetilde{f}_b(x) - K_b * f(x))^2 \, \mathrm{d}x > \delta^2 r^2/(72 U_n) \Big)$$

$$\lesssim U_n \delta^{-2} \sum_{b \in \mathcal{B}} \mathbb{E} \Big[\int_{\mathbb{R}} (\widetilde{f}_b(x) - K_b * f(x))^2 \, \mathrm{d}x \mathbb{1}_{B_\varepsilon(\min \mathcal{B})} \Big] + P\big((B_\varepsilon(\min \mathcal{B}))^c \big)$$

$$\lesssim \frac{(\log n)^2}{\delta^2 (n \wedge m) b^{2\beta+1}} + o(1). \qquad \qquad \square$$

Proof of Theorem 2.4. By a Taylor expansion argument we have

$$\widetilde{q}_{\tau,b} - q_\tau = -\frac{\widetilde{M}_b'(q_\tau)}{\widetilde{M}_b''(q_\tau^*)} = -\frac{\int_{-\infty}^{q_\tau} \widetilde{f}_b(x) \, \mathrm{d}x - \tau}{2 \widetilde{f}_b(q_\tau^*)} = -\frac{\int_{-\infty}^{q_\tau} (\widetilde{f}_b(x) - f(x)) \, \mathrm{d}x}{2 \widetilde{f}_b(q_\tau^*)},$$

for some intermediate point q_τ^* between q_τ and $\widetilde{q}_{\tau,b}$. By Proposition 2.2, the numerator in the above display is of order $\mathcal{O}_P(n^{-(\alpha+1)/(2\alpha+2\beta+1)})$ for the optimal bandwidth b^*. For the denominator we will show $\widetilde{f}_b(q_\tau^*) = f(q_\tau) + o_p(1)$ which completes the proof. Since $f(\bullet + q_\tau) \in C^\alpha([-\zeta, \zeta], R)$, we obtain $|f(x + q_\tau) - f(q_\tau)| < t/2$ for all $|x| \leqslant (\frac{t}{2R})^{1 \vee \alpha^{-1}} \wedge \zeta =: \delta$ for any $t > 0$. Therefore,

$$P(|\widetilde{f}_b(\widetilde{q}_{\tau,b}) - f(q_\tau)| > t) \leqslant P\big(\sup_{x \in [-\delta,\delta]} |\widetilde{f}_b(x + q_\tau) - f(q_\tau)| > t \big) + P(|\widetilde{q}_{\tau,b} - q_\tau| > \delta)$$

$$\leqslant P\big(\sup_{x \in [-\delta,\delta]} |\widetilde{f}_b(x + q_\tau) - f(x + q_\tau)| > t/2 \big) + P(|\widetilde{q}_{\tau,b} - q_\tau| > \delta). \tag{35}$$

Checking that the bandwidth satisfies $b \to 0$ and $\log(n)/(n b^{2\beta+1}) \to 0$ for $n \to \infty$, the first term on the right hand side above converges to zero by the uniform consistency proved in Proposition 2.3. The second one vanishes by Lemma A.4. $\qquad \square$

Appendix B: Proofs for Section 3

We start with showing that the construction of $\mathcal{B}_n$ from Lemma 3.1 satisfies Assumption 2.

Proof of Lemma 3.1. The deterministic counterpart of $\widetilde{b}_{min}$, which is defined in (12), is given by

$$b_{min} := \min\Big\{ b \in \Lambda_n : 2 \leqslant \Big(\frac{\log n}{n}\Big)^{1/2} \int_{-1/b}^{1/b} |\varphi_\varepsilon(u)|^{-1}\, du \leqslant 4 \Big\}.$$

Noting that for $f_\varepsilon \in \mathcal{D}^\beta(R,\gamma)$

$$4 \geqslant \Big(\frac{\log n}{n}\Big)^{1/2} \int_{-1/b_{min}}^{1/b_{min}} |\varphi_\varepsilon(u)|^{-1}\, du \gtrsim \Big(\frac{\log n}{n b_{min}^{2\beta+2}}\Big)^{1/2}$$

we obtain $n b_{min}^{2\beta+2} \to \infty$ and thus it is sufficient to prove the following

$$\inf_{f \in \mathcal{C}^\alpha(R,r,\zeta)} \inf_{f_\varepsilon \in \mathcal{D}^\beta(R,\gamma)} P(b_{min} \leqslant \widetilde{b}_{min} \leqslant b^*) \to 1 \quad \text{as} \quad n \to \infty. \tag{36}$$

for the optimal bandwidth $b^* = n^{-1/2\alpha+2(\beta \vee 1/2)+1}$ for convenience we define

$$I_n(b) := \Big(\frac{\log n}{n}\Big)^{1/2} \int_{-1/b}^{1/b} \frac{du}{|\varphi_\varepsilon(u)|}, \quad \widetilde{I}_n(b) := \Big(\frac{\log n}{n}\Big)^{1/2} \int_{-1/b}^{1/b} \frac{du}{|\varphi_{\varepsilon,m}(u)|}.$$

Assume $\widetilde{b}_{min} < b_{min}$, then monotonicity implies $\widetilde{I}_n(b_{min}) \leqslant \widetilde{I}_n(\widetilde{b}_{min}) \leqslant 1$. Combining with $I_n(b_{min}) \geqslant 2$, we obtain $I_n(b_{min}) - \widetilde{I}_n(b_{min}) \geqslant 1$. Hence,

$$P(\widetilde{b}_{min} < b_{min}) \leqslant P\big(|I_n(b_{min}) - \widetilde{I}_n(b_{min})| \geqslant 1\big). \tag{37}$$

On the other hand, if $b^* < \widetilde{b}_{min}$, we get $\widetilde{I}_n(b^*) \geqslant \widetilde{I}_n(\widetilde{b}_{min}) \geqslant 1/2$. Since $I_n(b^*) \lesssim (\frac{\log n}{n(b^*)^{2\beta+2}})^{1/2}$ converges to zero, $I_n(b^*) \leqslant 1/4$ for n large enough. Thus,

$$P(\widetilde{b}_{min} > b^*) \leqslant P\big(|I_n(b^*) - \widetilde{I}_n(b^*)| \geqslant 1/4\big). \tag{38}$$

To show that the right-hand sides of (37) and (38) converge to zero, we first apply the Cauchy–Schwarz inequality

$$\big|I_n(b) - \widetilde{I}_n(b)\big|^2 \leqslant \frac{\log n}{n} \int_{-1/b}^{1/b} \frac{du}{|\varphi_\varepsilon(u)|^2} \int_{-1/b}^{1/b} \Big|\frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)} - 1\Big|^2 du$$

$$\lesssim \frac{\log n}{n b^{2\beta+1}} \int_{-1/b}^{1/b} \Big|\frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)} - 1\Big|^2 du.$$

By Lemma A.1 we can consider the event $B_\varepsilon(b)$ only and thus Markov's inequality and (29) yield

$$P\Big(\big\{|I_n(b) - \widetilde{I}_n(b)| \geqslant \tfrac{1}{4}\big\} \cap B_\varepsilon(b)\Big) \lesssim \frac{\log n}{n b^{2\beta+1}} \int_{-1/b}^{1/b} \mathbb{E}\Big[\Big|\frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)} - 1\Big|^2 \mathbb{1}_{B_\varepsilon(b)}\Big]\, du \lesssim \frac{\log n}{n m b^{4\beta+2}}$$

which converges to zero for $b \in \{b_{min}, b^*\}$. Therefore, (36) holds true. $\qquad\square$

Before we can prove Theorem 3.2, some preparations are needed. By Lemma A.2 there is a constant $D > 0$ such that the bias can be bounded by $B_b := D b^{\alpha+1}$. By the error representation (8) we have for any $b \in \mathcal{B}$

$$|\widetilde{q}_{\tau,b} - q_\tau| = \Big|\frac{\int_{-\infty}^{q_\tau} (\widetilde{f}_b(x) - f(x))\, dx}{2\widetilde{f}_b(\widetilde{q}^*)}\Big| \leqslant \frac{B_b + |V_{b,X} + V_{b,\varepsilon} + V_{b,c}|}{2|\widetilde{f}_b(q^*)|} \tag{39}$$

with some $q^* \in [(q_\tau \wedge \widetilde{q}_{\tau,b}), (q_\tau \vee \widetilde{q}_{\tau,b})]$ and where the stochastic error is decomposed in

$$V_{b,X} := \frac{1}{n} \sum_{j=1}^n \left(\xi_j(b) - \mathbb{E}[\xi_j(b)] \right) \quad \text{with} \tag{40}$$

$$\xi_j(b) := \int_{-\infty}^0 a_s(x) \, \mathcal{F}^{-1} \left[\frac{\varphi_K(bu) e^{iuY_j}}{\varphi_\varepsilon(u)} \right] (x + q_\tau) \, \mathrm{d}x,$$

$$V_{b,\varepsilon} := \int_{-\infty}^0 a_s(x) \, \mathcal{F}^{-1} \left[\frac{\varphi_K(bu)\varphi_n(u)}{\varphi_\varepsilon(u)} \left(\frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)} - 1 \right) \right] (x + q_\tau) \, \mathrm{d}x \tag{41}$$

$$V_{b,c} := \int_{-\infty}^0 a_c(x) \, \mathcal{F}^{-1} \left[\varphi_K(bu) \left(\frac{\varphi_n(u)}{\varphi_{\varepsilon,m}(u)} - \varphi_X(u) \right) \right] (x + q_\tau) \, \mathrm{d}x. \tag{42}$$

In view of the analysis in Section A.2, the part of the stochastic error which is due to the continuous part a_c will be negligible. Hence, we concentrate on $V_{b,X}$ and $V_{b,\varepsilon}$. By independence of $(\xi_j(b))_j$, we obtain

$$\mathrm{Var}(V_{b,X}) \leqslant \frac{1}{n} \mathbb{E}[\xi_j(b)^2] = \frac{1}{n} \mathbb{E}\left[\left(\int_{-\infty}^0 a_s(x) \, \mathcal{F}^{-1} \left[\frac{\varphi_K(bu) e^{iuY_j}}{\varphi_\varepsilon(u)} \right] (x + q_\tau) \, \mathrm{d}x \right)^2 \right] =: \sigma_{b,X}^2. \tag{43}$$

To determine the variance of $V_{b,\varepsilon}$, it will be helpful again to restrict to the event $B_\varepsilon(b)$, defined in (18). We apply Plancherel's identity and the Cauchy–Schwarz inequality to separate Y_i and ε_i from each other since they are not necessarily independent:

$$\begin{aligned}
\mathbb{E}[|V_{b,\varepsilon}| \mathbb{1}_{B_\varepsilon(b)}] =& \frac{1}{2\pi} \mathbb{E}\left[\left| \int_{\mathbb{R}} \mathcal{F} a_s(-u) e^{-iuq_\tau} \frac{\varphi_K(bu)\varphi_n(u)}{\varphi_\varepsilon(u)} \left(\frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)} - 1 \right) \mathrm{d}u \right| \mathbb{1}_{B_\varepsilon(b)} \right] \\
\leqslant& \frac{1}{2\pi} \mathbb{E}\left[\left(\int_{\mathbb{R}} |\varphi_K(bu)| \left| \frac{\varphi_n(u)}{\varphi_\varepsilon(u)} \right|^2 \mathrm{d}u \right)^{1/2} \right. \\
& \times \left. \left(\int_{\mathbb{R}} |\varphi_K(bu)| |\mathcal{F} a_s(-u)|^2 \left| \frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)} - 1 \right|^2 \mathrm{d}u \right)^{1/2} \mathbb{1}_{B_\varepsilon(b)} \right] \\
\leqslant& \frac{1}{2\pi} \mathbb{E}\left[\left(\int_{\mathbb{R}} |\varphi_K(bu)| \left| \frac{\varphi_n(u)}{\varphi_\varepsilon(u)} \right|^2 \mathrm{d}u \right)^{1/2} \right. \\
& \times \left. \sup_{|u| \leqslant 1/b} |\varphi_{\varepsilon,m}(u) - \varphi_\varepsilon(u)| \left(\int_{\mathbb{R}} |\varphi_K(bu)| \left| \frac{\mathcal{F} a_s(-u)}{\varphi_{\varepsilon,m}(u)} \right|^2 \mathrm{d}u \mathbb{1}_{B_\varepsilon(b)} \right)^{1/2} \right]. \tag{44}
\end{aligned}$$

Let us define

$$\sigma_{b,\varepsilon} := \frac{1}{2\pi} m^{-1/2} \sigma_{b,\varepsilon,1} \sigma_{b,\varepsilon,2} \tag{45}$$

with

$$\sigma_{b,\varepsilon,1} := \mathbb{E}\left[\left(\int_{\mathbb{R}} |\varphi_K(bu)| \left| \frac{\varphi_n(u)}{\varphi_\varepsilon(u)} \right|^2 \mathrm{d}u \right)^{1/2} \right],$$

$$\sigma_{b,\varepsilon,2} := \mathbb{E}\left[\left(\int_{\mathbb{R}} |\varphi_K(bu)| \left| \frac{\mathcal{F} a_s(-u)}{\varphi_{\varepsilon,m}(u)} \right|^2 \mathrm{d}u \right)^{1/2} \mathbb{1}_{B_\varepsilon(b)} \right].$$

With the bounds $\sigma_{b,X}$ and $\sigma_{b,\varepsilon}$ at hand, we obtain the following concentration results.

Lemma B.1. *Let $\mathcal{B}$ be a set satisfying $|\mathcal{B}| \lesssim \log n$, $(\log \log n)/nb_1 \to 0$ for $b_1 = \min \mathcal{B}$ as well as $|\log b_1| \lesssim \log n$. Then we obtain uniformly over $f \in \mathcal{C}^\alpha(R, r, \zeta)$ and $f_\varepsilon \in \mathcal{D}^\beta(R, \gamma)$ for any $\delta > 0$:*

(i) $\quad P\left(\exists b \in \mathcal{B} : |V_{b,X}| \geqslant (1 + \delta) \sqrt{\log \log n} \left(\sqrt{2} \sigma_{b,X} + o(n^{-1/2}(b^{-\beta+1/2} \vee 1)) \right) \right) \to 0.$

(ii) $\quad P\left(\exists b \in \mathcal{B} : |V_{b,\varepsilon}| \geqslant \delta(\log n)^3 \sigma_{b,\varepsilon} \right) \to 0.$

(iii) *Assuming further $mb_1^{(2\beta \wedge 1)+2} \gtrsim 1$,*
$\quad P\left(\exists b \in \mathcal{B} : |V_{b,c}| \geqslant (\log n)^{3/2} n^{-1/2}(b^{-\beta+1/2} \vee 1) \right) \to 0.$

Proof. (i) Using Bernstein's inequality, we start with proving uniformly over $f \in \mathcal{C}^{\alpha}(R, r, \zeta)$ and $f_{\varepsilon} \in \mathcal{D}^{\beta}(R, \gamma)$ for any positive $\kappa_n = o(nb)$

$$P\big(|V_{b,X}| \geqslant \sqrt{\kappa_n}\big(\sqrt{2}\sigma_{b,X} + o(n^{-1/2}(b^{-\beta+1/2} \vee 1)))\big)\big) \leqslant 2e^{-\kappa_n}. \tag{46}$$

The assertion follows then from choosing $\kappa_n = (1+\delta)^2 \log \log n$ for some $\delta > 0$ and $|\mathcal{B}| \lesssim \log n$.

Plancherel's identity yields the deterministic bound

$$|\xi_j(b)| = \Big|\int_{-1/b}^{1/b} \mathcal{F}\, a_s(u) \frac{\varphi_K(bu)e^{iuY_j}}{\varphi_{\varepsilon}(u)} e^{-iuq_{\tau}}\, du\Big| \leqslant \int_{-1/b}^{1/b} |\mathcal{F}\, a_s(u)| \Big|\frac{\varphi_K(bu)}{\varphi_{\varepsilon}(u)}\Big|\, du$$

$$\lesssim \int_{-1/b}^{1/b} \frac{1}{(1+|u|)|\varphi_{\varepsilon}(u)|}\, du \lesssim b^{\beta}$$

Hence, $|\xi_j(b) - \mathbb{E}[\xi_j(b)]| \leqslant Cb^{-\beta}$ for some constant $C > 0$. Since the variance is bounded by (43), Bernstein's inequality (e.g. Massart (2007), Prop. 2.9) yields

$$P\Big(|V_{b,X}| \geqslant \sqrt{2\sigma_{b,X}^2 \kappa_n} + \frac{C\kappa_n}{3nb^{\beta}}\Big) \leqslant 2e^{-\kappa_n}.$$

Therefore, (46) follows from $\sqrt{\kappa_n}(nb^{\beta})^{-1} \lesssim (n(b^{2\beta-1} \wedge 1))^{-1/2}\big(\kappa_n/(nb)\big)^{1/2}$.

(ii) Using an estimate as in (44), we obtain

$$|V_{b,\varepsilon}| \leqslant \frac{1}{2\pi}\Big(\int_{\mathbb{R}} |\varphi_K(bu)|\Big|\frac{\varphi_n(u)}{\varphi_{\varepsilon}(u)}\Big|^2\, du\Big)^{1/2} \Big(\int_{\mathbb{R}} |\varphi_K(bu)||\mathcal{F}\, a_s(-u)|^2 \Big|\frac{\varphi_{\varepsilon}(u) - \varphi_{\varepsilon,m}(u)}{\varphi_{\varepsilon,m}(u)}\Big|^2\, du\Big)^{1/2}$$

$$\leqslant \frac{1}{2\pi}\Big(\int_{\mathbb{R}} |\varphi_K(bu)|\Big|\frac{\varphi_n(u)}{\varphi_{\varepsilon}(u)}\Big|^2\, du\Big)^{1/2}$$

$$\times \Big(\int_{\mathbb{R}} |\varphi_K(bu)|\Big|\frac{\mathcal{F}\, a_s(-u)}{\varphi_{\varepsilon,m}(u)}\Big|^2\, du\Big)^{1/2} \sup_{|u| \leqslant 1/b} \big|\varphi_{\varepsilon}(u) - \varphi_{\varepsilon,m}(u)\big|$$

$$=: \frac{1}{2\pi}V_{b,\varepsilon,1}V_{b,\varepsilon,2} \sup_{|u| \leqslant 1/b} \big|\varphi_{\varepsilon}(u) - \varphi_{\varepsilon,m}(u)\big|.$$

Hence for any $c > 0$

$$P\Big(\big\{|V_{b,\varepsilon}| \geqslant \delta(\log n)^3 \sigma_{b,\varepsilon}\big\} \cap B_{\varepsilon}(b)\Big)$$

$$\leqslant P\big(|V_{b,\varepsilon,1}| \geqslant (\log n)^{1+c}\sigma_{b,\varepsilon,1}\big) + P\Big(\big\{|V_{b,\varepsilon,2}| \geqslant (\log n)^{1+c}\sigma_{b,\varepsilon,2}\big\} \cap B_{\varepsilon}(b)\Big)$$

$$+ P\Big(\sup_{|u| \leqslant 1/b} \big|\varphi_{\varepsilon}(u) - \varphi_{\varepsilon,m}(u)\big| \geqslant \delta(\log n)^{1-2c}m^{-1/2}\Big) =: P_{b,1} + P_{b,2} + P_{b,3}.$$

The first two probabilities can be bounded by Markov's inequality:

$$P_{b,1} \leqslant (\log n)^{-1-c}\sigma_{b,\varepsilon,1}^{-1} \mathbb{E}[V_{b,\varepsilon,1}] = (\log n)^{-1-c},$$

$$P_{b,2} \leqslant (\log n)^{-1-c}\sigma_{b,\varepsilon,2}^{-1} \mathbb{E}[V_{b,\varepsilon,2}\mathbb{1}_{B_{\varepsilon}(b)}] = (\log n)^{-1-c}.$$

For $P_{b,3}$ we will apply the following version of Talagrand's inequality (cf. Massart (2007), (5.50)): Let T be a countable index and for all $t \in T$ let $Z_{1,t}, \ldots, Z_{n,t}$ be an i.i.d. sample of centered, complex valued random variables satisfying $\|Z_{k,t}\|_{\infty} \leqslant b$, for all $t \in T, k = 1, \ldots, n$, as well as $\sup_{t \in T} \mathrm{Var}(\sum_{k=1}^{n} Z_{k,t}) \leqslant v < \infty$. Then for all $\kappa > 0$

$$P\Big(\sup_{t \in T}\Big|\sum_{k=1}^{n} Z_{k,t}\Big| \geqslant 4\,\mathbb{E}\Big[\sup_{t \in T}\Big|\sum_{k=1}^{n} Z_{k,t}\Big|\Big] + \sqrt{2v\kappa} + \frac{2}{3}b\kappa\Big) \leqslant 2e^{-\kappa}. \tag{47}$$

Choosing the rational numbers $T = \mathbb{Q} \cap [-\frac{1}{b}, \frac{1}{b}]$ and $Z_{k,t} := e^{it\varepsilon_k^*} - \varphi_\varepsilon(t)$, Talagrand's inequality applies with $b = 2$ and $v = n$. As in (17) we use Theorem 4.1 by Neumann and Reiß (2009) to obtain for any $\eta \in (0, 1/2)$

$$m^{1/2} \mathbb{E}\left[\sup_{|u| \leqslant 1/b} |\varphi_{\varepsilon,m}(t) - \varphi_\varepsilon(t)|\right] \lesssim |\log b|^{1/2 + \eta}.$$

Therefore on the assumptions $\kappa_n^{-1}(\log n)^{1+2\eta} \to 0$ and $\kappa_n/m \to 0$

$$4\,\mathbb{E}\left[\sup_{|u| \leqslant 1/b, u \in \mathbb{Q}} |\varphi_{\varepsilon,m}(u) - \varphi_\varepsilon(u)|\right] + \sqrt{\frac{2\kappa_n}{m}} + \frac{4}{3m}\kappa_n = \sqrt{\frac{\kappa_n}{m}}\left(\sqrt{2} + o(1)\right)$$

and thus continuity of $\varphi_{\varepsilon,m}$ and (47) yield

$$P_{b,3} = P\left(\sup_{|u| \leqslant 1/b, u \in \mathbb{Q}} |\varphi_{\varepsilon,m}(u) - \varphi_\varepsilon(u)| \geqslant (\sqrt{2} + o(1))\sqrt{\kappa_n/m}\right) \leqslant 2e^{-\kappa_n}.$$

With $\kappa_n = \frac{\delta}{2}(\log n)^{2-4c}$ for $c < 1/4 - \eta/2$ we obtain $P_3 \leqslant 2n^{-\delta/2}$.

Using $b_1 = \min \mathcal{B}$, $|\log B| \lesssim \log n$ and Lemma A.1, we finally get

$$P\left(\sup_{b \in \mathcal{B}} |V_{b,\varepsilon}| \geqslant (\sqrt{2} + \delta)(\log n)^3 \sigma_{b,\varepsilon}\right) \leqslant \sum_{b \in \mathcal{B}} \left(P_{b,1} + P_{b,2} + P_{b,3}\right) + P(B_\varepsilon(b_1)^c) = o(1).$$

(iii) Corollary A.3 shows for $\delta_b > 0$ and for any sequence $(x_n)_n$ that tends to infinity

$$P\left(\exists b \in \mathcal{B} : |V_{b,c}| \geqslant \delta_b\right) \lesssim \sum_{b \in \mathcal{B}} \frac{x_n}{\delta_b^2 n(mb^{2\beta+2} \wedge 1)} + o(1).$$

So, choosing $\delta_b = (\log n)^{3/2} n^{-1/2}(b^{-\beta+1/2} \vee 1)$ and $x_n = o((\log n)^{1/2})$ yields

$$P\left(\exists b \in \mathcal{B} : |V_{b,c}| \geqslant (\log n)^{3/2} n^{-1/2}(b^{-\beta+1/2} \vee 1)\right)$$
$$\lesssim \sum_{b \in \mathcal{B}} \frac{x_n}{(\log n)^3 (mb^{(2\beta \wedge 1)+2} \wedge 1)} + o(1) \lesssim \frac{x_n}{(\log n)^2 (mb^{(2\beta \wedge 1)+2} \wedge 1)} + o(1) = o(1).$$

<div align="right">□</div>

For the denominator in the error representation (39) we need uniform consistency. A uniform result on the error $|\widetilde{q}_{\tau,b} - q_\tau|$ follows immediately.

Lemma B.2. *Let $\mathcal{B}$ be a finite set satisfying $|\mathcal{B}| \lesssim \log n$, $\sup_{b \in \mathcal{B}} b \log(n) \to 0$ as well as $\sup_{b \in \mathcal{B}}(\log n)^2/(nb^{2\beta+1}) \to 0$. Then we obtain for $n \to \infty$ and $\eta \in (0, 1)$*

$$\sup_{f \in \mathcal{C}^\alpha(R, r, \zeta)} \sup_{f_\varepsilon \in \mathcal{D}^\beta(R, \gamma)} P\left(\sup_{b \in \mathcal{B}} \sup_{q_\tau^* \in [q_\tau \wedge \widetilde{q}_{\tau,b}, q_\tau \vee \widetilde{q}_{\tau,b}]} |\widetilde{f}_b(q_\tau^*) - f(q_\tau)| > \eta f(q_\tau)\right) \to 0. \qquad (48)$$

Moreover, supposing $\min_{b \in \mathcal{B}} nb^{(2\beta \wedge 1)+2} \gtrsim 1$, we obtain uniformly in $f \in \mathcal{C}^\alpha(R, r, \zeta)$ and $f_\varepsilon \in \mathcal{D}^\beta(R, \gamma)$ for any sequence of critical values $(\delta_b)_{b \in \mathcal{B}}$ satisfying $\inf_{\mathcal{B}} \delta_b \to \infty$

$$P\left(\exists b \in \mathcal{B} : |\widetilde{q}_{\tau,b} - q_\tau| > \delta_b\left(3Db^{\alpha+1} + n^{-1/2}(b^{-\beta+1/2} \vee 1)\right)\right) \lesssim \sum_{b \in \mathcal{B}} \frac{1}{\delta_b} + o(1). \qquad (49)$$

Proof. Since $f(q_\tau) \geqslant r$ and $f \in C^\alpha([q_\tau - \zeta, q_\tau + \zeta], R)$, decomposition (35) implies with $\kappa = (\frac{\eta r}{2R})^{1 \vee \alpha^{-1}} \wedge \zeta$

$$P\left(\sup_{b \in \mathcal{B}} \sup_{q_\tau^* \in [q_\tau \wedge \widetilde{q}_{\tau,b}, q_\tau \vee \widetilde{q}_{\tau,b}]} |\widetilde{f}_b(q_\tau^*) - f(q_\tau)| > \eta f(q_\tau)\right)$$

$$\leqslant P\Big(\sup_{b\in\mathcal{B}}\sup_{x\in[-\kappa,\kappa]}|\widetilde{f}_b(x+q_\tau)-f(x+q_\tau)|>\eta r/2\Big)+P\Big(\sup_{b\in\mathcal{B}}|\widetilde{q}_{\tau,b}-q_b|>\kappa\Big). \tag{50}$$

Using $b_1=\min\mathcal{B}$, the first probability can be bounded by

$$\sum_{b\in\mathcal{B}}P\Big(\Big\{\sup_{x\in[-\kappa,\kappa]}|\widetilde{f}_b(x+q_\tau)-f(x+q_\tau)|>\eta r/2\Big\}\cap B_\varepsilon(b_1)\Big)+P(B_\varepsilon(b_1)^c)$$

$$\lesssim\log n\sup_{b\in\mathcal{B}}P\Big(\Big\{\sup_{x\in[-\kappa,\kappa]}|\widetilde{f}_b(x+q_\tau)-f(x+q_\tau)|>\eta r/2\Big\}\cap B_\varepsilon(b_1)\Big)+o(1)=o(1),$$

since for all b the probability in the last line converges faster to zero than $1/\log n$ owing to the concentration inequalities (31) and (32) and the conditions on b. To estimate the second term in (50), we apply Lemma A.4. Therefore, the conditions $b\log(n)\to 0$ and $(\log n)^2/(nb^{2\beta+1})\to 0$ yield the first assertion.

The estimate (49) follows from the error decomposition (8), (48) and Corollary A.3 with $x_n=o(\inf_\mathcal{B}\delta_b)$

$$P\Big(\exists b\in\mathcal{B}:|\widetilde{q}_{\tau,b}-q_\tau|>\delta_b\big(3Db^{\alpha+1}+n^{-1/2}(b^{-\beta+1/2}\vee 1)\big)\Big)$$

$$\leqslant P\Big(\exists b\in\mathcal{B}:\Big|\int_{-\infty}^{q_\tau}\widetilde{f}_b(x)-f(x)\,\mathrm{d}x\Big|>\tfrac{1}{2}f(q_\tau)\delta_b\big(3Db^{\alpha+1}+n^{-1/2}(b^{-\beta+1/2}\vee 1)\big)\Big)$$

$$+P\Big(\sup_{b\in\mathcal{B}}\sup_{q_\tau^*\in[q_\tau\wedge\widetilde{q}_{\tau,b},q_\tau\vee\widetilde{q}_{\tau,b}]}|\widetilde{f}_b(q_\tau^*)-f(q_\tau)|>\tfrac{1}{2}f(q_\tau)\Big)$$

$$\lesssim\sum_{b\in\mathcal{B}}\Big(\frac{1}{\delta_b}+\frac{1}{\delta_b^2}\frac{x_n}{mb^{1\wedge2\beta+2}\wedge 1}\Big)+o(1)\lesssim\sum_{b\in\mathcal{B}}\frac{1}{\delta_b}+o(1). \qquad\square$$

The variances $\sigma_{b,X}$ and $\sigma_{b,\varepsilon}$, defined in (43) and (45), are of course not available in practice. Instead, they can be estimated by

$$\widetilde{\sigma}_{b,X}^2=\frac{1}{n^2}\sum_{j=1}^n\Big(\int_{-\infty}^0 a_s(x)\,\mathcal{F}^{-1}\Big[\frac{\varphi_K(bu)e^{iuY_j}}{\varphi_{\varepsilon,m}(u)}\Big](x+\widetilde{q}_{\tau,b})\,\mathrm{d}x\Big)^2,\quad\widetilde{\sigma}_{b,\varepsilon}^2=\frac{1}{2\pi}m^{-1/2}\widetilde{\sigma}_{b,\varepsilon,1}^2\widetilde{\sigma}_{b,\varepsilon,2}^2$$

with

$$\widetilde{\sigma}_{b,\varepsilon,1}^2=\int_{-1/b}^{1/b}|\varphi_K(bu)|\Big|\frac{\varphi_n(u)}{\varphi_{\varepsilon,m}(u)}\Big|^2\,\mathrm{d}u,\quad\widetilde{\sigma}_{b,\varepsilon,2}^2=\int_{-1/b}^{1/b}|\varphi_K(bu)|\frac{|\mathcal{F}\,a_s(u)|^2}{|\varphi_{\varepsilon,m}|^2}\,\mathrm{d}u.$$

The following two lemmas show that these estimators are indeed reasonable.

Lemma B.3. *Let $\mathcal{B}$ be a finite set satisfying $|\mathcal{B}|\lesssim\log n$, $\max_{b\in\mathcal{B}}b^\alpha\log n\to 0$ as well as $\min_{b\in\mathcal{B}}nb^{2\beta+2}\to\infty$. Then we obtain for all $\eta>0$ as $n\to\infty$*

$$\sup_{f\in\mathcal{C}^\alpha(R,r,\zeta)}\sup_{f_\varepsilon\in\mathcal{D}^\beta(R,\gamma)}P\Big(\exists b\in\mathcal{B}:|\widetilde{\sigma}_{b,X}-\sigma_{b,X}|>\eta m^{-1/2}(b^{-\beta+1/2}\vee 1)\Big)\to 0.$$

Proof. Note that

$$\widetilde{\sigma}_{b,X}^2=\frac{1}{n^2}\sum_{j=1}^n\xi_{j,1}^2(b)+\frac{1}{n^2}\sum_{j=1}^n\xi_{j,2}^2(b)+\frac{1}{n^2}\sum_{j=1}^n\xi_{j,3}^2(b)$$

$$+\frac{2}{n^2}\sum_{j=1}^n\xi_{j,1}(b)\xi_{j,2}(b)+\frac{2}{n^2}\sum_{j=1}^n\xi_{j,1}(b)\xi_{j,3}(b)+\frac{2}{n^2}\sum_{j=1}^n\xi_{j,2}(b)\xi_{j,3}(b), \tag{51}$$

where we have defined

$$\xi_{j,1}(b):=\int_{-\infty}^0 a_s(x)\,\mathcal{F}^{-1}\Big[\varphi_K(bu)e^{iuY_j}\Big(\frac{1}{\varphi_{\varepsilon,m}(u)}-\frac{1}{\varphi_\varepsilon(u)}\Big)\Big](x+\widetilde{q}_{\tau,b})\,\mathrm{d}x,$$

$$\xi_{j,2}(b) := \int_{-\infty}^{0} a_s(x)\,\mathcal{F}^{-1}\left[\frac{\varphi_K(bu)e^{iuY_j}}{\varphi_\varepsilon(u)}\right](x+q_\tau)\,\mathrm{d}x,$$

$$\xi_{j,3}(b) := \int_{-\infty}^{0} a_s(x)\,\mathcal{F}^{-1}\left[\frac{\varphi_K(bu)e^{iuY_j}(e^{-iu\widetilde{q}_{\tau,b}}-e^{-iuq_\tau})}{\varphi_\varepsilon(u)}\right](x)\,\mathrm{d}x.$$

We will first study these three terms separately. Applying Plancherel's identity, the Cauchy–Schwarz inequality, the Neumann type bound (29) as well as $|\mathcal{F}\,a_s(u)| \leqslant A_s(1+|u|)^{-1}$ and the decay of φ_ε, we obtain

$$\mathbb{E}[|\xi_{j,1}(b)|^2 \mathbb{1}_{B_\varepsilon(b)}] \leqslant \frac{9}{2\pi^2}\int_{-1/b}^{1/b}\frac{|\mathcal{F}\,a_s(u)|^2}{|\varphi_\varepsilon(u)|^2}\,\mathrm{d}u\int_{-1/b}^{1/b}\frac{|\varphi_K(bu)|^2}{m|\varphi_\varepsilon(u)|^2}\,\mathrm{d}u \lesssim \frac{1}{(b^{2\beta-1}\wedge 1)mb^{2\beta+1}}, \quad (52)$$

$$\mathbb{E}[|\xi_{j,2}(b)|^2] = \mathbb{E}\left[\left|\frac{1}{2\pi}\int_{\mathbb{R}}\mathcal{F}\,a_s(u)e^{-iuq_\tau}\frac{\varphi_K(bu)}{\varphi_\varepsilon(u)}e^{iuY_j}\,\mathrm{d}u\right|^2\right]$$

$$\leqslant \frac{\|K\|_{L^1}^2 A_s^2 \|f_Y\|_\infty R^2}{4\pi^2}\int_{-1/b}^{1/b}(1+|u|)^{2\beta-2}\,\mathrm{d}u =: S_b^2 \quad (53)$$

as well as the deterministic bound

$$|\xi_{j,2}(b)|^2 = \left|\frac{1}{2\pi}\int_{\mathbb{R}}\mathcal{F}\,a_s(u)e^{-iuq_\tau}\frac{\varphi_K(bu)}{\varphi_\varepsilon(u)}e^{iuY_j}\,\mathrm{d}u\right|^2 \leqslant \frac{\|K\|_{L_1}^2 A_s^2}{4\pi^2}\int_{-1/b}^{1/b}(1+|u|)^{2\beta}\,\mathrm{d}u =: d_b^2. \quad (54)$$

Hence, $\mathrm{Var}[\xi_{j,2}(b)^2] \leqslant \mathbb{E}[\xi_{j,2}(b)^4] \leqslant d_b^2 S_b^2$. and $|\xi_{j,2}^2(b) - \mathbb{E}[\xi_{j,2}^2(b)]| \leqslant 2d_b^2$, so that an application of Bernstein's inequality yields for any $b > 0$ and $z > 0$

$$P\left(\left|\frac{1}{n}\sum_{j=1}^{n}(\xi_{j,2}^2(b) - \mathbb{E}[\xi_{j,2}^2(b)])\right| \geqslant z\right) \leqslant 2\exp\left(-\frac{z^2 n}{2S_b^2 d_b^2 + \frac{4}{3}d_b^2 z}\right).$$

Setting $z = S_b^2$ and noting $S_b^2 \lesssim (b^{-2\beta+1} \vee 1), d_b^2 \lesssim b^{-2\beta}$, we see that

$$P\left(\left|\frac{1}{n}\sum_{j=1}^{n}(\xi_{j,2}^2(b) - \mathbb{E}[\xi_{j,2}^2(b)])\right| \geqslant S_b^2\right) \leqslant 2\exp\left(-\frac{S_b^2 n}{4d_b^2}\right) \leqslant 2\exp\left(-Cnb^{2\beta\wedge 1}\right) \quad (55)$$

for some constant C. The right-hand side of (55) tends to zero with polynomial rate since $nb^{2\beta\wedge 1} \gtrsim \log n$.

Finally, we use $\mathrm{supp}\,a_s \subseteq [-1, 0]$ to write $\xi_{j,3}$ as

$$\xi_{j,3}(b) = \int_{\mathbb{R}}\left(a_s(x - \widetilde{q}_{\tau,b}) - a_s(x - q_\tau)\right)\mathcal{F}^{-1}\left[\frac{\varphi_K(bu)e^{iuY_j}}{\varphi_\varepsilon(u)}\right](x)\,\mathrm{d}x$$

$$\leqslant \sup_{t\in(-1,0)}|a_s'(t)||\widetilde{q}_{\tau,b} - q_\tau|\int_{(\widetilde{q}_{\tau,b}\wedge q_\tau)-1}^{\widetilde{q}_{\tau,b}\vee q_\tau}\mathcal{F}^{-1}\left[\frac{\varphi_K(bu)e^{iuY_j}}{\varphi_\varepsilon(u)}\right](x)\,\mathrm{d}x.$$

The Cauchy–Schwarz inequality and Plancherel's identity yield

$$|\xi_{j,3}(b)|^2 \leqslant \|a_s'\mathbb{1}_{(-1,0)}\|_\infty^2|\widetilde{q}_{\tau,b} - q_\tau|^2(1 + |\widetilde{q}_{\tau,b} - q_\tau|)\int_{(\widetilde{q}_{\tau,b}\wedge q_\tau)-1}^{\widetilde{q}_{\tau,b}\vee q_\tau}\left|\mathcal{F}^{-1}\left[\frac{\varphi_K(bu)e^{iuY_j}}{\varphi_\varepsilon(u)}\right](x)\right|^2\mathrm{d}x$$

$$\leqslant \frac{\|a_s'\mathbb{1}_{(-1,0)}\|_\infty^2}{2\pi}|\widetilde{q}_{\tau,b} - q_\tau|^2(1 + |\widetilde{q}_{\tau,b} - q_\tau|)\int_{\mathbb{R}}\left|\frac{\varphi_K(bu)}{\varphi_\varepsilon(u)}\right|^2\mathrm{d}u$$

$$\lesssim |\widetilde{q}_{\tau,b} - q_\tau|^2(1 + |\widetilde{q}_{\tau,b} - q_\tau|)b^{-2\beta-1}.$$

By Lemma A.4 $\sup_{b\in\mathcal{B}}|\widetilde{q}_{\tau,b} - q_\tau| = o_P(1)$. Applying (49), we conclude for some constant $C > 0$, for $\delta_b = (b^{\alpha+(1/2-\beta)_+} + n^{-1/2}b^{-\beta-1/2}))^{-1}$ and for any $\eta > 0$

$$P\left(\exists b \in \mathcal{B} : |\xi_{j,3}(b)| > \eta(b^{-\beta+1/2} \vee 1)\right)$$

$$\leqslant P\Big(\exists b \in \mathcal{B} : |\widetilde{q}_{\tau,b} - q_\tau| > \eta C b^{(\beta \wedge 1/2)+1/2}\Big) + o(1)$$

$$\leqslant P\Big(\exists b \in \mathcal{B} : |\widetilde{q}_{\tau,b} - q_\tau| > \eta C \delta_b(b^{\alpha+1} + n^{-1/2}(b^{-\beta+1/2} \vee 1))\Big) + o(1)$$

$$\lesssim \Big(\sum_{b \in \mathcal{B}} (\delta_b)^{-1}\Big) + o(1)$$

$$\lesssim \sup_{b \in \mathcal{B}} b^\alpha \log n + \sup_{b \in \mathcal{B}} \frac{\log n}{\sqrt{n} b^{\beta+1/2}} + o(1) = o(1). \tag{56}$$

Combining the variance bounds (52), (53) and (56), we apply Markov's inequality, the Cauchy–Schwarz inequality and the concentration result (55) on the decomposition (51) to obtain

$$\sup_{b \in \mathcal{B}} \Big(n(b^{2\beta-1} \wedge 1)|\widetilde{\sigma}_{b,X}^2 - \sigma_{b,X}^2|\Big) = \sup_{b \in \mathcal{B}} \Big(\frac{b^{2\beta-1} \wedge 1}{n} \sum_{j=1}^n (\xi_{j,2}^2(b) - \mathbb{E}[\xi_{j,2}^2(b)])\Big) + o_P(1) = o_P(1).$$

$\square$

Lemma B.4. *Let $\mathcal{B}$ be a finite set satisfying $|\mathcal{B}| \lesssim \log n$ as well as $\sup_{b \in \mathcal{B}} 1/(nb^{2\beta+1}) \to 0$. Then we obtain uniformly over $f \in \mathcal{C}^\alpha(R, r, \zeta)$ and $f_\varepsilon \in \mathcal{D}^\beta(R, \gamma)$ for all $\eta > 0$ as $n \to \infty$*

$$P\Big(\exists b \in \mathcal{B} : |\widetilde{\sigma}_{b,\varepsilon} - \sigma_{b,\varepsilon}| > \eta(\log n)m^{-1/2}(b^{-\beta+1/2} \vee 1)\Big) \to 0.$$

Proof. We start by showing for $b_1 = \min \mathcal{B}$ that

$$\sup_{|u| \leqslant 1/b_1} \Big|\frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)}\Big| = 1 + o_P(1). \tag{57}$$

To this end, recall $w(u) = (\log(e + |u|))^{-1/2-\eta}$ for some $\eta \in (0, 1/2)$. Markov's inequality, Lemma A.1 and Theorem 4.1 by Neumann and Reiß (2009) yield for any $\delta > 0$

$$P\Big(\sup_{|u| \leqslant 1/b_1} \Big|\frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)} - 1\Big| \geqslant \delta\Big) \leqslant P\Big(\sup_{|u| \leqslant 1/b_1} m^{1/2}|\varphi_\varepsilon(u) - \varphi_{\varepsilon,m}(u)| \geqslant \delta|\log b_1|\Big)$$

$$+ P\Big(\inf_{|u| \leqslant 1/b_1} |\varphi_{\varepsilon,m}(u)| \leqslant m^{-1/2}|\log b_1|\Big)$$

$$\leqslant (\delta|\log b_1|)^{-1} \mathbb{E}\Big[\sup_{|u| \leqslant 1/b_1} m^{1/2}|\varphi_\varepsilon(u) - \varphi_{\varepsilon,m}(u)|\Big] + o(1)$$

$$\leqslant \frac{1}{\delta|\log b_1|w(1/b_1)} \mathbb{E}\Big[\sup_{u \in \mathbb{R}} m^{1/2}w(u)|\varphi_\varepsilon(u) - \varphi_{\varepsilon,m}(u)|\Big] + o(1)$$

$$= o(1),$$

which is (57). Note that $[-1/b_1, 1/b_1]$ is the maximal interval for all $b \in \mathcal{B}$ and thus (57) holds uniformly in $\mathcal{B}$.

Now, we consider $\widetilde{\sigma}_{b,\varepsilon,1}$. The uniform consistency (57) implies

$$\widetilde{\sigma}_{b,\varepsilon,1}^2 = (1 + o_P(1)) \int_{\mathbb{R}} |\varphi_K(bu)| \Big|\frac{\varphi_n(u)}{\varphi_\varepsilon(u)}\Big|^2 \, du.$$

Chebyshev's inequality yields for all $\eta > 0$

$$P\Big(\sup_{b \in \mathcal{B}} \Big|\Big(\int_{\mathbb{R}} |\varphi_K(bu)| \frac{|\varphi_n(u)|^2}{|\varphi_\varepsilon(u)|^2} \, du\Big)^{1/2} - \mathbb{E}\Big[\Big(\int_{\mathbb{R}} |\varphi_K(bu)| \frac{|\varphi_n(u)|^2}{|\varphi_\varepsilon(u)|^2} \, du\Big)^{1/2}\Big]\Big| > \eta \log n\Big)$$

$$\leqslant (\eta \log n)^{-2} \sum_{b \in \mathcal{B}} \mathbb{E}\Big[\int_{\mathbb{R}} |\varphi_K(bu)| \frac{|\varphi_n(u)|^2}{|\varphi_\varepsilon(u)|^2} \, du\Big]$$

$$\lesssim (\eta^2 \log n)^{-1} \int_{-1/b_1}^{1/b_1} \frac{\mathbb{E}[|\varphi_n(u)|^2]}{|\varphi_\varepsilon(u)|^2} \, du \lesssim (\eta^2 \log n)^{-1},$$

where the last estimate follows from $\mathbb{E}[|\varphi_n(u)|^2] \lesssim |\varphi_Y(u)|^2 + \mathbb{E}[|\varphi_n(u) - \varphi_Y(u)|^2] \lesssim |\varphi_Y(u)|^2 + 1/n$, $f_\varepsilon \in \mathcal{D}^\beta(R, \gamma), \|f\|_\infty \lesssim 1$ and $nb_1^{2\beta+1} \to \infty$. Hence, we obtain uniformly in $\mathcal{B}$

$$\widetilde{\sigma}_{b,\varepsilon,1} = (1 + o_P(1))(\sigma_{b,\varepsilon,1} + o_P(\log n)) = \sigma_{b,\varepsilon,1} + o_P(\log n). \tag{58}$$

Concerning $\widetilde{\sigma}_{b,\varepsilon,2}$, we write with use of (57)

$$\widetilde{\sigma}_{b,\varepsilon,2}^2 = \int_{-1/b}^{1/b} |\varphi_K(bu)| \frac{|\mathcal{F} a_s(u)|^2}{|\varphi_{\varepsilon,m}(u)|^2} \, du = (1 + o_p(1)) \int_{-1/b}^{1/b} |\varphi_K(bu)| \frac{|\mathcal{F} a_s(u)|^2}{|\varphi_\varepsilon(u)|^2} \, du.$$

Moreover, the triangle inequality for the L^2-norm and Lemma A.1, applied on $B_\varepsilon(b_1)$ which is defined in (18), yield

$$\left| \left(\int_{-1/b}^{1/b} |\varphi_K(bu)| \frac{|\mathcal{F} a_s(u)|^2}{|\varphi_\varepsilon(u)|^2} \, du \right)^{1/2} - \sigma_{b,\varepsilon,2} \right|^2$$

$$\leqslant 2 \left| \mathbb{E} \left[\left(\left(\int_{-1/b}^{1/b} |\varphi_K(bu)| \frac{|\mathcal{F} a_s(u)|^2}{|\varphi_\varepsilon(u)|^2} \, du \right)^{1/2} - \left(\int_{-1/b}^{1/b} |\varphi_K(bu)| \frac{|\mathcal{F} a_s(u)|^2}{|\varphi_{\varepsilon,m}(u)|^2} \, du \right)^{1/2} \right) \mathbb{1}_{B_\varepsilon(b_1)} \right] \right|^2$$

$$+ 2P\big((B_\varepsilon(b_1))^c\big) \int_{-1/b}^{1/b} |\varphi_K(bu)| \frac{|\mathcal{F} a_s(u)|^2}{|\varphi_\varepsilon(u)|^2} \, du$$

$$\leqslant 2 \, \mathbb{E} \left[\left(\int_{-1/b}^{1/b} |\varphi_K(bu)| |\mathcal{F} a_s(u)|^2 \frac{|\varphi_{\varepsilon,m}(u) - \varphi_\varepsilon(u)|^2}{|\varphi_\varepsilon(u)\varphi_{\varepsilon,m}(u)|^2} \, du \right) \mathbb{1}_{B_\varepsilon(b_1)} \right] + o(1) \int_{-1/b}^{1/b} \frac{|\mathcal{F} a_s(u)|^2}{|\varphi_\varepsilon(u)|^2} \, du$$

$$\leqslant \frac{2}{|\log b_1|^{3/2}} \, \mathbb{E} \left[\int_{-1/b}^{1/b} \frac{|\mathcal{F} a_s(u)|^2}{|\varphi_\varepsilon(u)|^2} m |\varphi_{\varepsilon,m}(u) - \varphi_\varepsilon(u)|^2 \, du \right] + o(1)(b^{-2\beta+1} \vee 1)$$

$$= o(1)(b^{-2\beta+1} \vee 1),$$

where $o(1)$ is a null sequence which does not depend on b. Consequently,

$$\sup_{b \in \mathcal{B}} \left| \left(\int_{-1/b}^{1/b} |\varphi_K(bu)| \frac{|\mathcal{F} a_s(u)|^2}{|\varphi_\varepsilon(u)|^2} \, du \right)^{1/2} - \sigma_{b,\varepsilon,2} \right| (b^{\beta-1/2} \wedge 1) = o(1).$$

We get

$$\widetilde{\sigma}_{b,\varepsilon,2} = (1 + o_p(1))(\sigma_{b,\varepsilon,2} + o(b^{-\beta+1/2} \vee 1)) = \sigma_{b,\varepsilon,2} + o_P(b^{-\beta+1/2} \vee 1), \tag{59}$$

where the last estimate follows from $\sigma_{b,\varepsilon,2}^2 \lesssim b^{-2\beta+1} \vee 1$ by the analysis of the convergence rates.

Since $\sigma_{b,\varepsilon,1} \lesssim 1$ and $\sigma_{b,\varepsilon,2} \lesssim b^{-\beta+1/2} \vee 1$, it remains to combine (58) and (59) to obtain uniformly in $\mathcal{B}$

$$\widetilde{\sigma}_{b,\varepsilon} = \frac{1}{2\pi} m^{-1/2} \widetilde{\sigma}_{b,\varepsilon,1} \widetilde{\sigma}_{b,\varepsilon,2} = \frac{1}{2\pi} m^{-1/2} (\sigma_{b,\varepsilon,1} + o_P(\log n))(\sigma_{b,\varepsilon,2} + o_P(b^{-\beta+1/2} \vee 1))$$

$$= \sigma_{b,\varepsilon} + o_P\big((\log n) m^{-1/2}(b^{-\beta+1/2} \vee 1)\big). \qquad \square$$

Proof of Theorem 3.2. As seen in error decomposition (39), there are three stochastic errors $V_{b,X}, V_{b,\varepsilon}$ and $V_{b,c}$ which were treated in Lemma B.1. This motivates the following definition. For $\delta_1 > 0$ let

$$S_{b,X} := (1 + \delta_1) \sqrt{2 \log \log n} \max_{\mu \in \mathcal{B}: \mu \geqslant b} \sigma_{\mu,X}, \quad S_{b,\varepsilon} := (\delta_1 \log n)^3 \max_{\mu \in \mathcal{B}: \mu \geqslant b} \sigma_{\mu,\varepsilon}.$$

Note that on the assumption $|\varphi_\varepsilon(u)| \gtrsim (1+|u|)^{-\beta}$ we obtain for $\sigma_{b,\varepsilon} = \frac{1}{2\pi} m^{-1/2}\sigma_{b,\varepsilon,1}\sigma_{b,\varepsilon,2}$, defined in (45), that

$$\sigma_{b,\varepsilon,2}^2 \gtrsim \int_{-1/b}^{1/b} |\mathcal{F} a_s(-u)|^2 (1+|u|)^{2\beta} \, \mathrm{d}u \gtrsim \int_{-1/b}^{1/b} (1+|u|)^{2\beta-2} \, \mathrm{d}u \sim b^{-2\beta+1} \vee 1$$

and thus $\sigma_{b,\varepsilon} \gtrsim m^{-1/2}(b^{-\beta+1/2} \vee 1)$. Therefore, Lemma B.1 yields

$$P\Big(\exists b \in \mathcal{B} : |V_{b,X} + V_{b,\varepsilon} + V_{b,c}| \geqslant S_{b,X} + S_{b,\varepsilon}\Big)$$

$$\leqslant P\Big(\exists b \in \mathcal{B} : |V_{b,X}| \geqslant S_{b,X} + \tfrac{1}{3}S_{b,\varepsilon}\Big) + P\Big(\exists b \in \mathcal{B} : |V_{b,\varepsilon}| \geqslant \frac{S_{b,\varepsilon}}{3}\Big) + P\Big(\exists b \in \mathcal{B} : |V_{b,c}| \geqslant \frac{S_{b,\varepsilon}}{3}\Big)$$

$$=o(1).$$

Hence, the probability of the event

$$A_1 := \Big\{\forall b \in \mathcal{B} : |V_{b,X} + V_{b,\varepsilon} + V_{b,c}| \leqslant S_{b,X} + S_{b,\varepsilon}\Big\}$$

converges to one. The variances $S_{b,X}$ and $S_{b,\varepsilon}$ can be estimated by

$$\widetilde{S}_{b,X} := (1+\delta_1)\sqrt{2\log\log n} \max_{\mu \in \mathcal{B}:\mu \geqslant b} \widetilde{\sigma}_{\mu,X}, \quad \widetilde{S}_{b,\varepsilon} := (\delta_1 \log n)^3 \max_{\mu \in \mathcal{B}:\mu \geqslant b} \widetilde{\sigma}_{\mu,\varepsilon}.$$

We apply Lemmas B.3 and B.4 and the triangle inequality of the ℓ^∞-norm to obtain uniformly in $b \in \mathcal{B}$

$$\big|\max_{\mu \geqslant b} \widetilde{\sigma}_{\mu,X} - \max_{\mu \geqslant b} \sigma_{\mu,X}\big| \leqslant \max_{\mu \geqslant b} |\widetilde{\sigma}_{\mu,X} - \sigma_{\mu,X}| = o_P\Big(\frac{1}{m^{1/2}(b^{\beta-1/2} \wedge 1)}\Big),$$

$$\big|\max_{\mu \geqslant b} \widetilde{\sigma}_{\mu,\varepsilon} - \max_{\mu \geqslant b} \sigma_{\mu,\varepsilon}\big| \leqslant \max_{\mu \geqslant b} |\widetilde{\sigma}_{\mu,\varepsilon} - \sigma_{\mu,\varepsilon}| = o_P\Big(\frac{\log n}{m^{1/2}(b^{\beta-1/2} \wedge 1)}\Big).$$

Using again $\sigma_{b,\varepsilon} \gtrsim m^{-1/2}(b^{-\beta+1/2} \vee 1)$, we thus obtain for all $\eta > 0$ that the event

$$A_2 := \Big\{\forall b \in \mathcal{B} : \big|(\widetilde{S}_{b,X} + \widetilde{S}_{b,\varepsilon}) - (S_{b,X} + S_{b,\varepsilon})\big| \leqslant \eta(S_{b,X} + S_{b,\varepsilon})\Big\}$$

has asymptotically probability one. The same holds true for the events

$$A_3 := \Big\{\forall b \in \mathcal{B} : \sup_{q^* \in [(q_\tau \wedge \widetilde{q}_{\tau,b}) \vee (q_\tau \wedge \widetilde{q}_{\tau,b})]} |\widetilde{f}_b(q^*) - f(q_\tau)| \leqslant \eta f(q_\tau)\Big\},$$

$$A_4 := \Big\{\forall b \in \mathcal{B} : \sup_{q^* \in [(q_\tau \wedge \widetilde{q}_{\tau,b}) \vee (q_\tau \wedge \widetilde{q}_{\tau,b})]} |\widetilde{f}_b(q^*) - \widetilde{f}_b(\widetilde{q}_{\tau,b})| \leqslant \eta|\widetilde{f}_b(\widetilde{q}_{\tau,b})|\Big\}$$

by (48). Therefore, it is sufficient to work in the following on the event

$$A := A_1 \cap A_2 \cap A_3 \cap A_4.$$

We now show that the adaptive estimator $\widetilde{q}_\tau$ mimics the oracle estimator defined as follows. Recalling the estimate of the bias $B_b = Db^{\alpha+1}$, let the oracle bandwidth be defined by

$$b_* := \max\{b \in \mathcal{B} : B_b \leqslant S_{b,X} + S_{b,\varepsilon}\}. \tag{60}$$

Note that b_* is well-defined and unique since B_b is monoton increasing in b while $(S_{b,X} + S_{b,\varepsilon})$ is monton decreasing. We get the oracle estimator $\widetilde{q}_{\tau,b_*}$.

Since on A_4 for all $b \in \mathcal{B}$ and $q^* \in [(q_\tau \wedge \widetilde{q}_{\tau,b}) \vee (q_\tau \wedge \widetilde{q}_{\tau,b})]$

$$|\widetilde{f}_b(q^*)| \geqslant |\widetilde{f}_b(\widetilde{q}_{\tau,b})| - |\widetilde{f}_b(q^*) - \widetilde{f}_b(\widetilde{q}_{\tau,b})| \geqslant (1-\eta)|\widetilde{f}_b(\widetilde{q}_{\tau,b})|,$$

we have for any $b \in \mathcal{B}$ on the event $A_1 \cap A_4$ by (39)

$$|\widetilde{q}_{\tau,b} - q_\tau| \leqslant \frac{B_b + |V_{b,X} + V_{b,\varepsilon} + V_{b,c}|}{2|\widetilde{f}_b(q^*)|} \leqslant \frac{B_b + S_{b,X} + S_{b,\varepsilon}}{2(1-\eta)|\widetilde{f}_b(\widetilde{q}_{\tau,b})|}.$$

Furthermore, by the definition of b_* we have on the event A for any $b \leqslant b_*$

$$|\widetilde{q}_{\tau,b} - q_\tau| \leqslant \frac{S_{b,X} + S_{b,\varepsilon}}{(1-\eta)|\widetilde{f}_b(\widetilde{q}_{\tau,b})|}.$$

On A_2 we can estimate $\widetilde{S}_{b,X} + \widetilde{S}_{b,\varepsilon} \geqslant (1-\eta)(S_{b,X} + S_{b,\varepsilon})$. In particular, we have on the event A for any $b \leqslant b_*$

$$|\widetilde{q}_{\tau,b} - q_\tau| \leqslant \frac{\widetilde{S}_{b,X} + \widetilde{S}_{b,\varepsilon}}{(1-\eta)^2|\widetilde{f}_b(\widetilde{q}_{\tau,b})|}$$

Since for any $\delta > 0$ we find $\delta_1, \eta > 0$ such that $\left((1-\eta)^{-2}(\sqrt{2} + \delta_1) - \sqrt{2}\right) \vee \left((1-\eta)^{-2}\delta_1\right) < \delta$, we obtain $|\widetilde{q}_{\tau,b} - q_\tau| \leqslant \widetilde{\Sigma}_b$ with $\widetilde{\Sigma}_b$ as defined in (13). As a result one has $q_\tau \in \mathcal{U}_b$ and $q_\tau \in \mathcal{U}_\mu$ for all $b \leqslant b_*$ and $\mu \leqslant b_*$, implying $\mathcal{U}_\mu \cap \mathcal{U}_b \neq \emptyset$. By the definition of the procedure, $\widetilde{b}^* \geqslant b_*$ and $\mathcal{U}_{\widetilde{b}_*} \cap \mathcal{U}_{b_*} \neq \emptyset$ on the event A. This leads to

$$|\widetilde{q}_{\tau,\widetilde{b}^*} - q_\tau| \leqslant |\widetilde{q}_{\tau,b_*} - q_\tau| + |\widetilde{q}_{\tau,\widetilde{b}^*} - \widetilde{q}_{\tau,b_*}| \leqslant \widetilde{\Sigma}_{b_*} + (\widetilde{\Sigma}_{b_*} + \widetilde{\Sigma}_{\widetilde{b}^*})$$

On $A_2 \cap A_3$ we have $\widetilde{\Sigma}_b \lesssim S_{b,X} + S_{b,\varepsilon}$ since $f(q_\tau) \geqslant r$. Using additionally the monotonicity of $(S_{b,X} + S_{b,\varepsilon})$ as well as $\widetilde{b}^* \geqslant b_*$, this implies

$$|\widetilde{q}_{\tau,\widetilde{b}^*} - q_\tau| \lesssim (S_{b_*,X} + S_{b_*,\varepsilon}) \lesssim \left(\sqrt{\log\log n} + \left(\log n^\delta\right)^3\right)\left(b_*^{-\beta+1/2} \vee 1\right)n^{-1/2}.$$

It remains to note by the definition (60) of the oracle b_* and by the assumption $b_{j+1}/b_j \lesssim 1$ that $b_* \sim n^{-1/(2\alpha+2(\beta\vee 1/2)+1)}$ as $n \to \infty$. $\qquad\square$

References

CARROLL, R. J. and HALL, P. (1988). Optimal rates of convergence for deconvolving a density. *Journal of the American Statistical Association* **83** 1184–1186.

CARROLL, R. J., RUPPERT, D., STEFANSKI, L. A. and CRAINICEANU, C. M. (2006). *Measurement error in nonlinear models: a modern perspective* **105**. Chapman & Hall/CRC.

COMTE, F. and LACOUR, C. (2011). Data-driven density estimation in the presence of additive noise with unknown distribution. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **73** 601–627.

DATTNER, I., GOLDENSHLUGER, A. and JUDITSKY, A. (2011). On deconvolution of distribution functions. *Ann. Statist.* **39** 2477–2501.

DATTNER, I. and REISER, B. (2013). Estimation of distribution functions in measurement error models. *J. Statist. Plann. Inference* **143** 479–493.

DELAIGLE, A., HALL, P. and MEISTER, A. (2008). On deconvolution with repeated measurements. *Ann. Statist.* **36** 665–685.

DUDLEY, R. M. (1992). Fréchet Differentiability, p-Variation and Uniform Donsker Classes. *The Annals of Probability* **20** 1968–1982.

FAN, J. (1991). On the Optimal Rates of Convergence for Nonparametric Deconvolution Problems. *Ann. Statist.* **19** 1257-1272.

FRESE, E. M., FICK, A. and SADOWSKY, H. S. (2011). Blood pressure measurement guidelines for physical therapists. *Cardiopulmonary physical therapy journal* **22** 5.

GIRARDI, M. and WEIS, L. (2003). Operator–valued Fourier multiplier theorems on Besov spaces. *Mathematische Nachrichten* **251** 34–51.

GOLDENSHLUGER, A. and NEMIROVSKI, A. (1997). On spatially adaptive estimation of nonparametric regression. *Mathematical Methods of Statistics* **6** 135–170.

HALL, P. and LAHIRI, S. N. (2008). Estimation of distributions, moments and quantiles in deconvolution problems. *Ann. Statist.* **36** 2110–2134.

JOHANNES, J. (2009). Deconvolution with unknown error distribution. *Ann. Statist.* **37** 2301–2323.

JOHANNES, J. and SCHWARZ, M. (2010). Adaptive circular deconvolution by model selection under unknown error distribution. *arXiv:0912.1207v2*.

KANNEL, W. et al. (1995). Framingham study insights into hypertensive risk of cardiovascular disease. *Hypertension research: official journal of the Japanese Society of Hypertension* **18** 181–196.

KAPPUS, J. (2012). Nonparametric adaptive estimation for discretely observed Lévy processes. Available at http://edoc.hu-berlin.de/dissertationen/kappus-julia-johanna-2012-09-28/PDF/kappus.pdf.

LEPSKIĬ, O. V. (1990). A problem of adaptive estimation in Gaussian white noise. *Teor. Veroyatnost. i Primenen.* **35** 459–470.

LOUNICI, K. and NICKL, R. (2011). Global uniform risk bounds for wavelet deconvolution estimators. *Ann. Statist.* **39** 201–231.

MASSART, P. (2007). *Concentration inequalities and model selection. Lecture Notes in Mathematics* **1896**. Springer, Berlin.

MCMURRY, T. L. and POLITIS, D. N. (2004). Nonparametric regression with infinite order flat-top kernels. *Journal of Nonparametric Statistics* **16** 549–562.

MEISTER, A. (2004). On the effect of misspecifying the error density in a deconvolution problem. *Canadian Journal of Statistics* **32** 439–449.

NEUMANN, M. H. (1997). On the effect of estimating the error density in nonparametric deconvolution. *Journal of Nonparametric Statistics* **7** 307–330.

NEUMANN, M. H. (2007). Deconvolution from panel data with unknown error distribution. *Journal of Multivariate Analysis* **98** 1955–1968.

NEUMANN, M. H. and REISS, M. (2009). Nonparametric estimation for Lévy processes from low-frequency observations. *Bernoulli* **15** 223–248.

NICKL, R. and REISS, M. (2012). A Donsker theorem for Lévy measures. *J. Funct. Anal.* **263** 3306–3332.

ROSENTHAL, H. P. (1970). On the subspaces of L^p ($p > 2$) spanned by sequences of independent random variables. *Israel J. Math.* **8** 273–303.

SÖHL, J. and TRABS, M. (2012). A uniform central limit theorem and efficiency for deconvolution estimators. *Electron. J. Statist.* **6** 2486-2518.

SPOKOINY, V. and VIAL, C. (2009). Parameter tuning in pointwise adaptation using a propagation approach. *Ann. Statist.* **37** 2783–2807.

STIRNEMANN, J. J., COMTE, F. and SAMSON, A. (2012). Density estimation of a biomedical variable subject to measurement error using an auxiliary set of replicate observations. *Statistics in Medicine*.

TRIEBEL, H. (2010). *Theory of Function Spaces*. Birkhäuser Verlag, Basel Reprint of the 1983 Edition.

VAN DER VAART, A. W. (1998). *Asymptotic Statistics*. Cambridge University Press, Cambridge.